

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.6
(without alignments)

AW845712 MRO-CT006
AU10149 Drosophili
B182340 60304151
A1800420 tJ4b10_x
A2981841 2M0262123
BB311470 60142205
BG931641 AGENCOURT
BB929882 R05-GN004
A740764 nz03c06_nz03c06
AQ495774 HS_S219_B
AQ77664 CTR-HSP-2
BH069321 nz29g07_s
BG481055 60229036
BG788930 SEAMC008
AG093150 Pan trogl
A1051406 ox40h11_s
AQ077664 CTR-HSP-2
AQ495774 HS_S219_B
BH069339 R0C1-2A-3
BG104689 MRE-FN020
BB723456 AV697956
AW358199 42179 MAR
AA807714 AA807714
AV720699 AV720699
BB104689 MRE-FN020
BB723456 AV697956
AW358199 42179 MAR
AA807714 AA807714
AV720699 AV720699
BG108181 60228048
BB165435 60177326
BG424440 60226941
BB19389 TCBR1P98
BR353355 QV1-HT063
AU106E06P AL206615 Tetraodon
BG108181 60228048
BB165435 60177326
BG424440 60226941
BB19389 TCBR1P98
BR353355 QV1-HT063
AW852891 UT-HF-BN0
W42502 zc8d04_s1
A2288367 RPCI-23-4
N90521 zB41b08_s1
BB903653 601675142
BB622030 BB622030
BQ066068 AGBNCOURT

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 03:34:33 ; Search time 1115 Seconds

4575.407 Million cell updates/sec

Title: US-09-955-807-1_COPY_59_373
Perfect score: 315
Sequence: 1 atgtgggttattctgagcc.....atgtgtcttcagecagccac 315

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: em_lestba:*

2: em_lesthium:*

3: em_estin:*

4: em_estmu:*

5: em_lestov:*

6: em_lestpl:*

7: em_estro:*

8: em_htc:*

9: qb_esti:*

10: qb_htc:*

11: qb_htc:*

12: qb_est3:*

13: qb_est4:*

14: qb_est5:*

15: em_estrun:*

16: em_lestrom:*

17: qb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pIn:*

21: em_gss_vrt:*

22: em_gss_fur:*

23: em_gss_man:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_ror:*

ALIGNMENTS

RESULT 1

BG024363

LOCUS BG024363

DEFINITION 602274356FL NIH-MGC_85 Homo sapiens

ACCESSION BG024363

VERSION BG024363.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Buterilia; Primates; Catarhini; Hominidae; Homo.

REFERENCE I (bases 1 to 746)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgbps1@mail.nih.gov

TISSUE Procurement: Louis Staudt, M.D., Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL).

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LILNL at:

http://image.llnl.gov

Plate: L1A10006 row: P column: 14

High quality sequence start: 2

High quality sequence stop: 540.

FEATURES Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

VERSION	BK030759.1	GI:	16744329
KEYWORDS	EST.		
SOURCE	cow.		
ORGANISM	Bos taurus		
REFERENCE	I (bases 1 to 389)	1	(bases 1 to 648)
AUTHORS	Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-Mckown, C.G., Partea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and Keele, J.W.	Authors	Dias Neto, E., Garcia Correa, R., Verjovskiy-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zogo, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bataa, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
COMMENT	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle	COMMENT	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
TITLE	Single Pass sequencing: Bases called and alt_trimmed with phred	JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL	v0.980504.e. Vector identified by cross_match with the -minscore 18	MEDLINE	20202663
MEDLINE	and -minmatch 12 options.	COMMENT	Contact: Simpson A.J.G.
FOR PRIMERS	FORWARD: AGCAACACGACTAGTACGGACGT	FEATURES	Laboratory of Cancer Genetics
BACKWARD: GTTTCCAGCCACGCTTACGGACG	source	source	Ludwig Institute for Cancer Research
Plate: 121 row: P column: 22	Location/Qualifiers	source	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Seq primer: ATTTGGTGACGACTTACG	1..389	source	Tel: +55-11-2707001
Query Match	Best local similarity 54.6%; Pred. No. 4.1; Length 389; matches 71; conservative 0; Mismatches 59; Indels 0; Gaps 0;	source	Fax: +55-11-2707001
Qy	65 ACCCGCCCTTCATGCGCCTTGAGTACTGCTGAGTGTAGCATGCCACTCCA 124	source	Email: astimpson@ludwig.org.br
Db	140 ACCCTGCTGTCATGCTGCCCTGACCGTACCAAGGTAGCAGTAGACGCCAGTCTT 81	source	This sequence was derived from the RPPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tt1=&t2=MRO:CT0069-07109-002&t3=1999-10-07&t4=1)
Qy	125 AGCATGGCTTGAACAATTCCTCTGCCCTCACACAGTACAGGATTACTGGCG 184	source	Seq primer: puc 18 forward
Db	80 GGCTTCCTTTCATCCTCCCTCCCTCTTCAGGTGCCTGCAGATGCCCTGG 21	source	High quality sequence start: 32
Qy	185 ACTCGGGGSC 194	source	High quality sequence stop: 585
Db	20 GCAGGGGSC 11	source	Location/Qualifiers
FEATURES	source	source	Location/Qualifiers
BASE COUNT	91 a 111 c 128 g 59 t	source	1..648
ORIGIN	Query Match	source	/organism="Homo sapiens"
BASE COUNT	165 a 157 c 164 g	source	/db_xref="taxon:6606"
ORIGIN	Query Match	source	/clone.lib="CT0069"
BASE COUNT	162 t	source	/dev_stage="adult"
ORIGIN	Query Match	source	/note="Organ: colon; Vector: pUC18; Site_1: Smal; Site_2:
BASE COUNT	648	source	Smal; A mini-library was made by cloning products derived from OPRTES PCR (U.S. Letters Patent Application No. 196 ,716 - Ludwig Institute for Cancer Research) profiles into PUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
ORIGIN	Query Match	source	/note="Organelle: Mitochondrion; Vector: pUC18; Site_1: EcoRI; Site_2: SalI; Insert length: 6606 bp; Strands: Both; Sequencing strategy: Sanger; Sequence type: cDNA; Sample: Clonal library; Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."
BASE COUNT	91 a 111 c 128 g 59 t	source	Best Local Similarity 55.8%; Pred. No. 8.8; Matches 87; Conservative 0; Mismatches 67; Indels 2; Gaps 1;
ORIGIN	Query Match	source	Best Local Similarity 55.8%; Pred. No. 8.8; Matches 87; Conservative 0; Mismatches 67; Indels 2; Gaps 1;
BASE COUNT	165 a 157 c 164 g	source	Qy 148 CCTGGCCCTTCACCAAGCACTGAGCATTCACTGGGGACTCGGGCAAAGGAGATCTG 207
ORIGIN	Query Match	source	Qy 148 CCTGGCCCTTCACCAAGCACTGAGCATTCACTGGGGACTCGGGCAAAGGAGATCTG 207
BASE COUNT	162 t	source	Db 613 CCTGGCTCTCTGGACAGTGCCTGCTTGGAGGGTGCCTCCAGAGAACCTGCTCC 554
ORIGIN	Query Match	source	Db 613 CCTGGCTCTCTGGACAGTGCCTGCTTGGAGGGTGCCTCCAGAGAACCTGCTCC 554
BASE COUNT	648	source	Qy 208 TCACAGACAAAGGAGCCAGGSCACACATGGCTTACCTCT -GCCCTGCTCACCCAT 265
ORIGIN	Query Match	source	Qy 208 TCACAGACAAAGGAGCCAGGSCACACATGGCTTACCTCT -GCCCTGCTCACCCAT 265
BASE COUNT	91 a 111 c 128 g 59 t	source	Db 553 ACCAGGTCACTTCAGAGAACACTGTGCCTCATGCTGCTTCATGCTGCTCACACCA 494
ORIGIN	Query Match	source	Qy 266 GGCTACATGGTGGACACTGAGAACATGTTG 301
BASE COUNT	162 t	source	Db 493 TGCTTGTGAGGCTTGACTGCTCAGGAATCTCTGT 458
ORIGIN	Query Match	source	RESULT 8
BASE COUNT	91 a 111 c 128 g 59 t	source	CNS012E7/c
ORIGIN	Query Match	source	LOCUS CNS012E7/c
DEFINITION	AW845712	source	1088 bp DNA linear GSS 26-JUL-1999
ACCESSION	MR0-CP0069-071099-002-a04	source	droso_melano_genom survey sequence T7 end of BAC
VERSION	AW845712.1	source	BCN07008 of DrosBAC Library from <i>Drosophila melanogaster</i> (fruit fly), genomic survey sequence.
KEYWORDS	EST.	source	All0149
SOURCE	human.	source	All0149.1 GI:5613060
ORGANISM	Homo sapiens	source	GSS.
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	ORGANISM	Drosophila melanogaster.
		ORGANISM	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidea; Drosophilidae; Drosophila.

/note="Organ: kidney; Vector: pORB7; Site_1: XbaI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XbaI sites using the following 5' adaptor: GGCAGAGG(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT	ORIGIN
110 a	(Stratagene)
217 c	California, Berkeley
187 g	using ZAP-cDNA synthesis kit
161 t	(Life Technologies)

Query Match 10.9%; Score 34.2; DB 10; Length 675;
Best Local Similarity 52.4%; Pred. No. 14; Mismatches 0;
Matches 98; Conservative 0; Indels 1; Gaps 1;

Qy 127 CATGGCTTGAACAACTGTCCTGCCCTCACCGCAGTGACAGATTCACTGGGAC 186
Db 658 CAGGCTGTCCAAAGCAACTGTTGGTAGGACACAA-CAGGCGGCCAAAGACATGGTCTCA 540

Qy 187 TCGGGGCAAGGAGACTGTGACAGAACAGACAAGAGGAGCAGGTACACATGGTGTAC 306
Db 598 CCATTCCTCCAAAGCAACTGTTGGTAGGACACAA-CAGGCGGCCAAAGACATGGTCTCA 540

Qy 247 CTGGCCCTGCCCACCCATGGTACAGATGGGTGGACACCTGAGCTCA 306
Db 539 CAGTGACAGCCTGCCAAATGCTGAAGTGGGTGAGACTACCGCAGTAGTCGGGAGCTCTCA 480

Qy 307 GCGAGCC 313
Db 479 GGGGCC 473

RESULT 13

BQ931641/c
LOCUS BQ931641 1300 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT_8807646 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6427035
ACCESSION BQ931641
VERSION 5
KEYWORDS mRNA sequence.
SOURCE EST.
ORGANISM human.

REFERENCE
1 (bases 1 to 1300)
AUTHORS NIH_MGC http://mgc.ncbi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov

Tissue Procurement: ARCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL url:
<http://image.llnl.gov>

Place: LUCM2611 row: g column: 04

High quality sequence start: 121
High quality sequence stop: 343.

FEATURES

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6427035"
/clone_id="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/note="Organ: lung; Vector: pORB7; Site_1: EcoRI; Site_2: XbaI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XbaI sites using the following 5' adaptor: GGCAGAGG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BERKELEY) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 263 a
ORIGIN 477 c
Match 10.9%; Score 34.2; DB 14; Length 1300;
Best Local Similarity 51.7%; Pred. No. 17; Mismatches 0;
Matches 78; Conservative 0; Indels 73; Gaps 0;

Qy 140 ATATGCCCTGCCCTCCACAGCAGTACAGGATTCAGGGACTCGGGGCAAGG 199
Db 1133 AAAGTAATCCGGTATGGCCAGGAGAACAAAGGGAGGGCAAGG 1074

Qy 200 AGACTGTCACAGAACAAAGAGGAGCAGGGCACACATGGTACCTCCGCCRCCTC 259
Db 1073 TGTTGTGTCGGCGGAGGTTGAAGAACACAGGGGGGTGTTGCAATTCCGGGRC 1014

Qy 260 ACCATGGTACATGGCTGGACACCTCAG 290
Db 1013 ACAAGAGGGGGTGTAGGTGTACATTTCGG 983

RESULT 14

BB929882
LOCUS BB929882 351 bp mRNA linear EST 02-OCT-2000
DEFINITION RC5-GN0046-290800-033-c10 GN0046 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE929882
VERSION BE929882.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Cetartiodactyla; Hominoidea; Homo.
REFERENCE I (bases 1 to 351)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjorski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., de Oliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 2020663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?1=&t2=RC5-GN0046-290800-033-c10&t3=2000-08-29&t4=1>)
Seq primer: pac 18 forward
High quality sequence start: 18
High quality sequence stop: 94.
Location/Qualifiers
1. .351
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="GN0046"
/dev_stage="Adult"
/note="Organ: placenta normal; Vector: puc18; Site_1: Small; Site_2: Small; A mini-library was made by cloning products derived from ORESTES PCR U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

	BASE COUNT	83 a	81 c	103 g	84 t	Matches
ORIGIN						103; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
Query Match		10.8%		Score 34; DB 9; Length 413;		QY
Best Local Similarity		59.2%		Pred. No. 12; Pred. 0;		Db
Matches	58;	Conservative 0;	Mismatches 40;	Indels 0;	Gaps 0;	Db
DEFINITION	AA740764/c	AA740764	nz03cc06.s1	NCI-CGAP_GCB1	Homo sapiens	linear mRNA sequence.
ACCESSION	AY740764	AY740764	AY740764.1	GI:2779356		EST.
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
Homo sapiens						
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. I (bases 1 to 413)						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
COMMENT						
CONTACT						
CDNA Library Preparation:						
Email: ccrapbs@email.nih.gov						
Tissue Procurement:						
Ph.D., Gerald Marti, M.D.						
CDNA Library Arrayed by:						
Greg Lennon, Ph.D.						
DNA Sequencing by:						
Washington University Genome Sequencing Center						
Clone distribution:						
NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLM at: www.bio.lnl.nih.gov/borl/image/image.html						
Insert Length:	1908					
Std Error:	0					
Seq Primer:	-40m13 fwd. ET from Amersham					
High quality sequence stop:	396.					
Location/Qualifiers						
source						
1. 413						
/organism="Homo sapiens"						
/db_xref="taxon:9606"						
/clone="IMAGE:1286698"						
/clone_id="NCI-CGAP_GCB1"						
/tissue_type="terminal center B cell"						
/lab_host="DH10B"						
/note="Vector: pMT2D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-, provided by Dr. Louis M. Staudt (NCI), Dr. David Allman (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer [5'-TGTTCACAATCTGAATGGGAGCCGCCTTTTTTTTT-3', 1]. Double stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pMT3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."						
BASE COUNT	139	a	102	c	116	g
ORIGIN						
Query Match		10.8%		Score 34; DB 9;		Length 413;
Best Local Similarity		47.2%		Pred. No. 13;		

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 03:55:38 ; Search time 112 Seconds
 (without alignments)
 4127.139 Million cell updates/sec

title: US-09-955-807-1_COPY_59_373

Perfect score: 315

Sequence: 1 atgcgggttattctgagcc.....atgtgtttcagegagccac 315

Scoring table: IDENTITY_NUC

Searched: Gapext 1.0

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pupna/us07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pupna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pupna/us06_PUBCOMB.seq:*

4: /cgn2_6/ptodata/2/pupna/us06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pupna/us07_PUBCOMB.seq:*

6: /cgn2_6/ptodata/2/pupna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pupna/us08_PUBCOMB.seq:*

8: /cgn2_6/ptodata/2/pupna/us08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pupna/us09_NEW_PUB.seq:*

10: /cgn2_6/ptodata/2/pupna/us09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pupna/us10_NEW_PUB.seq:*

12: /cgn2_6/ptodata/2/pupna/us10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pupna/us60_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pupna/us60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	315	100.0	1692	10 US-09-955-807-1
2	313.4	99.5	1200	10 US-09-955-807-13
3	304	96.5	384	10 US-09-955-807-5
4	36.2	11.5	3736	9 US-09-965-830-10
C	33.8	10.7	414	9 US-10-123-155-418
C	6	10.6	98865	10 US-09-710-609A-3
C	7	10.3	81826	9 US-10-175-523-197
C	8	32.4	822	10 US-09-893-737-293
C	9	32.4	10.3	3428 9 US-09-822-846-204
C	10	10.0	3012	9 US-10-151-54A-18
C	11	31	9.8	2762 9 US-10-128-072-13
C	12	31	9.8	2762 9 US-10-121-040-13
C	13	31	9.8	2762 9 US-10-123-90-13
C	14	31	9.8	2762 9 US-10-140-470-13
C	15	31	9.8	2762 9 US-10-175-746-13
C	16	31	9.8	2762 9 US-10-176-921-13
C	17	31	9.8	2762 9 US-10-176-921-13
C	18	31	9.8	2762 9 US-10-137-865-13
C	31	9.8	2762 9 US-10-140-474-13	

ALIGNMENTS

RESULT 1
 US-09-955-807-1

; Sequence 1, Application US/09955807
 ; Patent No. US20020132996A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lok, Si
 ; APPLICANT: Sheppard, Paul O.
 ; APPLICANT: Kindsvoegel, Wayne
 ; APPLICANT: Bort, Susan J.
 ; TITLE OF INVENTION: Secretory Protein-48
 ; FILE REFERENCE: 98-17C1
 ; CURRENT APPLICATION NUMBER: US/09/955,807
 ; CURRENT FILING DATE: 2001-09-19
 ; PRIORITY NUMBER: 60/102,679
 ; PRIORITY FILING DATE: 1998-10-01
 ; PRIORITY NUMBER: 09/410,603
 ; PRIORITY FILING DATE: 1999-10-01
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 1
 LENGTH: 1692
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (59)...(373)

US-09-955-807-1

Query Match Best local Similarity 100.0%; Pred. No. 4 2e-96; Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTGGGTATTCGTGAGCCCATGCCATGCCATGCCACTGGCCCTTCTCTTGGC 60

Db 59 ATGGCTGGGTATTCGTGAGCCCATGCCATGCCACTGGCCCTTCTCTTGGC 118

QY 61 CTACACCTGCCCCCTTCCTTGCCCTTGAGCTGCGAGCTGAGCCACT 120

Db 119 CTACACCTGCCCCCTTCCTTGCCCTTGAGCTGCGAGCTGAGCCACT 178

QY 121 CCCAACGCAAGGCGCTTGGAACTATGCTCTCTGCCCCCTCACCAACGAGCTGAGCAAGGATCACT 180

RESULT 2
US-09-955-807-13
; Sequence 13, Application US/09955807
; Patent No. US20020132996A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; PRIORITY FILING DATE: 1998-09-19
; PRIORITY APPLICATION NUMBER: 60/102,679
; PRIORITY FILING DATE: 1998-10-01
; PRIORITY APPLICATION NUMBER: 09/410,603
; NUMBER OF SEQ ID NOS: 17
; LENGTH: 384
; SEQ ID NO: 6
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)..(384)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: misc-feature
; LOCATION: (1)..(384)
; OTHER INFORMATION: n = A,T,C or G
; US-09-955-807-6
; Query Match 96 %; Score 304; DB 10; Length 384;
; Best Local Similarity 97.5%; Pred. No. 1.4e-92; 8; Indels 0; Gaps 0;
; Matches 307; Conservative 0; Mismatches 0;
; Query 1 ATGCCTGGTTATTCGAGCCATGCCATGGCACACCCACTTGCCCTCTCCCTTAGGC 60
; Db 59 ATGCCTGGTTATTCGAGCCATGCCATGGCACACCCACTTGCCCTCTCCCTTAGGC 118
; Db 61 CTACACCCCTGCCCTCTTGCCCTTGAGTTACTGTGCTGGAGTGTACTGTGCTGGAGTGTAGGCCCAC 120
; Db 119 CTACACCCCTGCCCTCTTGCCCTTGAGTTACTGTGCTGGAGTGTAGGCCCAC 178
; Query Match 99.5%; Score 313 4; DB 10; Length 12001;
; Best Local Similarity 99.7%; Pred. No. 2.5e-95; 0; Mismatches 314; Conservative 0; Indels 0; Gaps 0;
; Matches 314; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
; Query 1 ATGCCTGGTTATTCGAGCCATGCCATGGCACACCCACTTGCCCTCTCCCTTAGGC 60
; Db 1025B ATGCCTGGTTATTCGAGCCATGCCATGGCACACCCACTTGCCCTCTCCCTTAGGC 10317
; Query 61 CTACACCCCTGCCCTCTTGCCCTTGAGTTACTGTGCTGGAGTGTACTGTGCTGGAGTGTAGGCCCAC 120
; Db 1031B CTACACCCCTGCCCTCTTGCCCTTGAGTTACTGTGCTGGAGTGTAGGCCCAC 10377
; Query 121 CCCAAAGCATGGCTGGACAGATGTCCTCGCCCTCCACCAAGCAGTGTGAGGTTACTGTGCTGGAGTGTAGGCCCAC 180
; Db 1037B CCCAAAGCATGGCTGGACAGATGTCCTCGCCCTCCACCAAGCAGTGTGAGGTTACTGTGCTGGAGTGTAGGCCCAC 10437
; Query 181 GGGGACTCTGGGGCAAGAGGAGCTGTGACAGAACAGGACAGGTGACACATGG 240
; Db 1043B GGGGACTCTGGGGCAAGAGGAGCTGTGACAGAACAGGACAGGTGACACATGG 10497
; Query 241 TGTACCTCGCCCTGCCTGACCATGGTGAACCTCGAATCATGTG 300
; Db 1049B TGTACCTCGCCCTGCCTGACCATGGTGAACCTCGAATCATGTG 10557
; Query 301 TCTTCAGGCCAC 315
; Db 1055B TCTTCAGGCCAC 10572
; RESULT 3
; Sequence 3, Application US/09955807
; Patent No. US20020132996A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955,807
; PRIORITY FILING DATE: 2001-09-19
; PRIORITY APPLICATION NUMBER: 09/410,603
; NUMBER OF SEQ ID NOS: 17
; LENGTH: 384
; SEQ ID NO: 6
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)..(384)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: misc-feature
; LOCATION: (1)..(384)
; OTHER INFORMATION: n = A,T,C or G
; US-09-955-807-6
; Query Match 96 %; Score 304; DB 10; Length 384;
; Best Local Similarity 97.5%; Pred. No. 1.4e-92; 8; Indels 0; Gaps 0;
; Matches 307; Conservative 0; Mismatches 0;
; Query 1 ATGCCTGGTTATTCGAGCCATGCCATGGCACACCCACTTGCCCTCTCCCTTAGGC 60
; Db 59 ATGCCTGGTTATTCGAGCCATGCCATGGCACACCCACTTGCCCTCTCCCTTAGGC 118
; Db 61 CTACACCCCTGCCCTCTTGCCCTTGAGTTACTGTGCTGGAGTGTACTGTGCTGGAGTGTAGGCCCAC 120
; Db 119 CTACACCCCTGCCCTCTTGCCCTTGAGTTACTGTGCTGGAGTGTAGGCCCAC 178
; Query 121 CCCAAAGCATGGCTGGACAGATGTCCTCGCCCTCCACCAAGCAGTGTGAGGTTACTGTGCTGGAGTGTAGGCCCAC 180
; Db 239 GGGGACTCTGGGGCAAGAGGAGCTGTGACAGAACAGGACAGGTGACACATGG 298
; Query 241 TGTACCTCGCCCTGCCTGACCATGGTGAACCTCGAATCATGTG 300
; Db 299 TGTACCTCGCCCTGCCTGACCATGGTGAACCTCGAATCATGTG 358
; Query 301 TCTTCAGGCCAC 315
; Db 359 TCTTCAGGCCAC 373
; RESULT 4
; Sequence 4, Application US/09955807
; Patent No. US20020177201A1
; GENERAL INFORMATION:
; APPLICANT: Yamamoto Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: A novel potassium channel protein
; FILE REFERENCE: Y9903-PCT
; CURRENT APPLICATION NUMBER: US/09/965,830
; CURRENT FILING DATE: 2001-10-01
; PRIORITY APPLICATION NUMBER: 09/600,776
; PRIORITY FILING DATE: 2001-07-21
; PRIORITY APPLICATION NUMBER: JP P1998-346198

PRIOR FILING DATE: 1998-12-04
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 10
 LENGTH: 3736
 TYPE: DNA
 ORGANISM: Rattus sp.
 US-09-965-830-10

Query Match 11.5%; Score 36.2; DB 9; Length 3736;
 Best Local Similarity 48.8%; Pred. No. 0.043; Mismatches 0; Indels 0; Gaps 0;

Matches 98; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 13 TCTGAGGCCATGCCATGTGACACCCACTTGCCCTCTCCCTAGGCCAACCTGGCC 72
 Db 3132 TCTGACCCCAAGCTCTGTGAGCACAGCTCAGTCAGCTGGTCAGCACATCCAGAAC 3191

QY 73 CTTTCTCTGGCCCTCTGTAGTAACTGTTGCTGAGGTGATGACGCCACTCCAAGCATGGC 132

Db 3192 CTGGCCTTGGGCCAAGCCCTGCTGGGGCTTAATTACCTGGCATCCAGGGAGGAGC 3251

QY 133 CTGGAAACAATGCTCTGCCTCACAGCAGTACAGGTTCACTGGACTGGGG 192

Db 3252 TGGGCCTCTGGCCCTCTGGCTGAGCTGGTCAGCAGTCAGCTGGCTGG 3311

QY 193 GCAAAGGAGACTGTGTCAAA 213

Db 3312 ATTCTCTGACTTTAACAA 3332

RESULT 5
 US-10-123-155-418/c
 Sequence 418, Application US/10123155
 Publication No. US20030068794A1

GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tunes, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zenin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

FILE REFERENCE: P333091C30
 CURRENT APPLICATION NUMBER: US/10/123,155
 CURRENT FILING DATE: 2002-04-15

PRIOR APPLICATION removed - see Palm or File Wrapper
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO: 418
 LENGTH: 414
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-123-155-418

Query Match 10.7%; Score 33.8; DB 9; Length 414;
 Best Local Similarity 10.7%; Pred. No. 0.15; Mismatches 148; Indels 0; Gaps 0;

Matches 28; Conservative 85; Mismatches 148; Indels 0; Gaps 0;

QY 46 CTCCTCTCTAGGGCTACACCTGGCCCTCTGGTAGTACCTGTGGCTGGA 105
 Db 295 CTY.T.YST..SADNTR.Y.K.BT.MYR...N...MSY.TS.YB.HB..MARM 236

QY 106 GTGATGAGGCCACTCCGAAAGCATGGCCTGGACAAATGTCCTCCCTGCCACAGCA 165
 Db 235 .MBMK.DYS.MCRYW MN.A.DTSTDND.B.Y.Y.HYR.TCYD..SCS.B.KYMMNA 176
 QY 166 GNGACAGATTCAGTGGGACTCAGGGGCAAAGGAGACTGTGTCAAGACAAAGGAGC 225
 Db 175 GBM.BRG.R.TCDYHHP..DS..SN..HH.TY.MA..T.SSNTG.CY.BSY.T..YC.H 116
 QY 226 CAGGCGCACATGCTGCTACCCATGCTGACATGGGTGGACAC 285
 Db 115 .SYTHY.SWRRCYBS..KSC.T...A.MY..HS.NBCTS..MH..HBHSHTA..YH 56
 55 H.C.A...CAD.YS.SAD.R. 35

RESULT 6
 US-09-770-689A-3/c
 Sequence 3, Application US/09770689A

Patent No. US20020115171A1

GENERAL INFORMATION:
 APPLICANT: YAN, Chunhua et al.

TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS, NUCLEAR ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE

TITLE OF INVENTION: PROTEINS, AND USES THEREOF

FILE REFERENCE: C1001079

CURRENT APPLICATION NUMBER: US/09/770,689A

CURRENT FILING DATE: 2001-01-29

NUMBER OF SEQ ID NOS: 5
 SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO: 3
 LENGTH: 9885
 TYPE: DNA
 ORGANISM: HUMAN
 US-09-770-689A-3

Query Match 10.6%; Score 33.4; DB 10; Length 98865;
 Best Local Similarity 53.4%; Pred. No. 0.93; Mismatches 61; Indels 0; Gaps 0;

Matches 70; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 175 TMCATGGGAAATGGGGCAAGGAGACTGTTGACAAAGACAAAAGGAGGCCAGGGTCAC 234
 Db 7422 TCCCTGTGAGAGGGCAGCCAGTGTCTGTACTCAAAGACTGTGGAAACCAAGGAAG 7363

QY 235 ACATGGGTACCCCTGGCTGCCTACCCATGGCATGGGTGGACACTCTGAAAT 294

Db 7362 AGGCCCTGGACCCAAGCCATGCTCAGCCACTACTGTGAACTGTGCAAACTTGAAGT 7303

QY 295 CATGGGTCTC 305

Db 7302 CAGGTTCCTC 7292

RESULT 7
 US-10-175-523-197

Sequence 197, Application US/10175523
 Publication No. US2003009626A1

GENERAL INFORMATION:
 APPLICANT: Brockman, Jeffrey

APPLICANT: Evans, David

APPLICANT: Hook, Derek

APPLICANT: Klimczak, Leszek

APPLICANT: Laeng, Pascal

APPLICANT: Palfrayman, Michael

APPLICANT: Rajan, Prithi
 TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)

FILE REFERENCE: 3235/L0795-US3
 CURRENT APPLICATION NUMBER: US/10/175,523

CURRENT FILING DATE: 2002-06-18
 PRIOR APPLICATION NUMBER: US 60/299,151
 PRIOR FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: US 60/317,828

PRIOR APPLICATION NUMBER: US 60/325, 150
 PRIOR FILING DATE: 2001-03-25
 PRIORITY NUMBER: US 60/333, 047
 PRIORITY FILING DATE: 2001-11-14
 PRIORITY APPLICATION NUMBER: US 60/349, 936
 PRIORITY FILING DATE: 2002-01-18
 PRIORITY APPLICATION NUMBER: US 60/361, 834
 PRIORITY FILING DATE: 2002-03-04
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 197
 LENGTH: 81826
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)..(81826)
 OTHER INFORMATION: where n may be a or g or c or t/u, unknown, or other
 S-10-175-5-197

RESULT 8
 S-09-893-737-293/c
 Sequence 243, Application US/09893737
 Patent No. US20020110855A1
 GENERAL INFORMATION:
 APPLICANT: Sheppard, Paul O.
 APPLICANT: Presnell, Scott R.
 TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
 FILE REFERENCE: 00-41
 CURRENT APPLICATION NUMBER: US/09/893, 737
 CURRENT FILING DATE: 2001-06-28
 PRIORITY APPLICATION NUMBER: US 60/215, 446
 PRIORITY FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 329
 SOFTWARE: FastaSQ for Windows Version 3.0
 SEQ ID NO: 293
 LENGTH: 822
 TYPE: DNA
 ORGANISM: Homo sapiens
 NAME/KEY: CDS
 LOCATION: (1)..(822)
 S-09-893-737-293

RESULT 8
 S-09-893-737-293/c
 Sequence 243, Application US/09893737
 Patent No. US20020110855A1
 GENERAL INFORMATION:
 APPLICANT: Sheppard, Paul O.
 APPLICANT: Presnell, Scott R.
 TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
 FILE REFERENCE: 00-41
 CURRENT APPLICATION NUMBER: US/09/893, 737
 CURRENT FILING DATE: 2001-06-28
 PRIORITY APPLICATION NUMBER: US 60/215, 446
 PRIORITY FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 329
 SOFTWARE: FastaSQ for Windows Version 3.0
 SEQ ID NO: 293
 LENGTH: 822
 TYPE: DNA
 ORGANISM: Homo sapiens
 NAME/KEY: CDS
 LOCATION: (1)..(822)

RESULT 8
 S-09-893-737-293/c
 Sequence 243, Application US/09893737
 Patent No. US20020110855A1
 GENERAL INFORMATION:
 APPLICANT: Sheppard, Paul O.
 APPLICANT: Presnell, Scott R.
 TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
 FILE REFERENCE: 00-41
 CURRENT APPLICATION NUMBER: US/09/893, 737
 CURRENT FILING DATE: 2001-06-28
 PRIORITY APPLICATION NUMBER: US 60/215, 446
 PRIORITY FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 329
 SOFTWARE: FastaSQ for Windows Version 3.0
 SEQ ID NO: 293
 LENGTH: 822
 TYPE: DNA
 ORGANISM: Homo sapiens
 NAME/KEY: CDS
 LOCATION: (1)..(822)

RESULT 8
 S-09-893-737-293/c
 Sequence 243, Application US/09893737
 Patent No. US20020110855A1
 GENERAL INFORMATION:
 APPLICANT: Sheppard, Paul O.
 APPLICANT: Presnell, Scott R.
 TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
 FILE REFERENCE: 00-41
 CURRENT APPLICATION NUMBER: US/09/893, 737
 CURRENT FILING DATE: 2001-06-28
 PRIORITY APPLICATION NUMBER: US 60/215, 446
 PRIORITY FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 329
 SOFTWARE: FastaSQ for Windows Version 3.0
 SEQ ID NO: 293
 LENGTH: 822
 TYPE: DNA
 ORGANISM: Homo sapiens
 NAME/KEY: CDS
 LOCATION: (1)..(822)

RESULT 8
 S-09-893-737-293/c
 Sequence 243, Application US/09893737
 Patent No. US20020110855A1
 GENERAL INFORMATION:
 APPLICANT: Sheppard, Paul O.
 APPLICANT: Presnell, Scott R.
 TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
 FILE REFERENCE: 00-41
 CURRENT APPLICATION NUMBER: US/09/893, 737
 CURRENT FILING DATE: 2001-06-28
 PRIORITY APPLICATION NUMBER: US 60/215, 446
 PRIORITY FILING DATE: 2000-06-30
 NUMBER OF SEQ ID NOS: 329
 SOFTWARE: FastaSQ for Windows Version 3.0
 SEQ ID NO: 293
 LENGTH: 822
 TYPE: DNA
 ORGANISM: Homo sapiens
 NAME/KEY: CDS
 LOCATION: (1)..(822)

RESULT 9
 US-09-822-846-204
 Sequence 204, Application US/09822846
 Publication No. US20030027139A1
 GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 APPLICANT: McCoy, John M.
 APPLICANT: Lavallie, Edward R.
 APPLICANT: Collins-Racie, Lisa A.
 APPLICANT: Evans, Cheryl
 APPLICANT: Merberg, David
 APPLICANT: Teacy, Maurice
 APPLICANT: Agostino, Michael J.
 APPLICANT: Steininger II, Robert J.
 APPLICANT: Bowman, Michael R.
 APPLICANT: Spaulding, Vicki
 APPLICANT: Wong, Gordon G.
 APPLICANT: Clark, Hilary
 APPLICANT: Fechtel, Kim
 APPLICANT: Howes, Steven H.
 APPLICANT: Resnick, Richard J.
 APPLICANT: Gulukota, Kamalakar
 APPLICANT: Graham, James R.
 APPLICANT: Genetics Institute, Inc.
 TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
 FILE REFERENCE: GIN 6400
 CURRENT APPLICATION NUMBER: US/09/822, 846
 CURRENT FILING DATE: 2001-03-29
 PRIORITY APPLICATION NUMBER: 60/195, 605
 PRIORITY FILING DATE: 2000-04-06
 NUMBER OF SEQ ID NOS: 629
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 204
 LENGTH: 3428
 TYPE: DNA
 ORGANISM: Homo sapiens
 NAME/KEY: CDS
 LOCATION: (1)..(3428)

RESULT 10
 US-10-151-542A-18/c
 Sequence 18, Application US/10151542A
 Publication No. US20030096348A1
 GENERAL INFORMATION:
 APPLICANT: Chen, Fang
 TITLE OF INVENTION: DNA MOLECULES ENCODING MAMMALIAN NUCLEAR
 RECEPTOR PROTEIN, DNRS
 FILE REFERENCE: 20083P1CA
 CURRENT APPLICATION NUMBER: US/10/151, 542A
 CURRENT FILING DATE: 2002-05-20
 PRIORITY APPLICATION NUMBER: 09/326, 755

PRIOR FILING DATE: 1999-06-04
PRIOR APPLICATION NUMBER: 09/581,033
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: PCT/US98/26422
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 60/069,379
PRIOR FILING DATE: 1997-12-12
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 18
LENGTH: 3012
TYPE: DNA
ORGANISM: Homo sapien (human)
US-10-151-542A-18

Query Match 10.0%; Score 31.4; DB 9; Length 3012;
Best Local Similarity 49.1%; Pred. No. 1,7; Mismatches 86; Indels 0; Gaps 0;
Matches 83; Conservative 0; MisMatches 86; Indels 0; Gaps 0;

QY 63 ACACCCCTGGCCCTTGCTTGCCCTTGAGTACTGTGCGCTGGATGAGTCACCTCC 122
Db 1247 AGACCCCTGGCCCTGTAGGCCTCATCATCCTCTAGGTGTCACCCGACCAT 1188

QY 123 CAAGCAGTGGCCCTGGAACAATGTCCTCTGCCCTCCACAGCAGTGGAGATTCACTG 182
Db 1187 CATCCCCATGCCAGATGCAAGCACCCACCGTACCATCCTAGGTGTCACCCGACCAT 1128

QY 183 GRACTGGGGCAAGGAGACGTGTCAGAACAGACAAGGAGCAGGGT 231
Db 1127 CAGCGCTGAACTAGAGTATGTCGTCGCCAGCAACCCVCAAGCT 1079

RESULT 11
US-10-028-072-1/C
; Sequence 13, Application US/10028072
; Publication No. US20030004311A1

GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurner, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watansabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang

TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063551
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064448
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065346
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09

PRIOR APPLICATION NUMBER: 60/077791
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085379
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349

Query Match 9 8%; Score 31; DB 9; Length 2762;
Query Local Similarity 62.0%; Pred. No. 2.2; Matches 49; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Query 80 TGCCTCTGTTGATCTGACTGCGCTGAGTCATGGCCACTCCACAGCATGGCTGGAAC 139
Db 1973 TCCCCACTCTAGTAGAGGGCTGGGGTATGGAGACCGTCCAGCCGGCGCTGCAGC 1914
Query 140 ATGGCTCTCTGCCCTCC 158
Db 1913 CTCTCTACTCCCTGGGCCTCC 1895

RESULT 12
US-10-121-049-13/C
Sequence 13, Application US10121049
Publication No. US20030022239A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beasini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filavoroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watansabe, Colin K.
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330P1C17
CURRENT APPLICATION NUMBER: US10/121,049
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 13
LENGTH: 2762
TYPE: DNA
ORGANISM: Homo Sapien
US-10-121-049-13

Mon Jun 23 10:42:32 2003

us-09-955-807-1_copy_59_373.rnpb

Search completed: June 21, 2003, 04:55:35
Job time : 114 secs

OM nucleic - nucleic search, using sw model
Run on: June 21, 2003, 03:27:58 ; Search time 39 Seconds
Title: US-09-955-807-1_COPY_59_373
Perfect score: 315
Sequence: 1 atgcgtggtttattctgagcc.....atgtgttttcacgcgaccc 315
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:
1: /cgn2_6/podata/1/ina/5A_COMB.seq:*
2: /cgn2_6/podata/1/ina/5B_COMB.seq:*
3: /cgn2_6/podata/1/ina/6A_COMB.seq:*
4: /cgn2_6/podata/1/ina/6B_COMB.seq:*
5: /cgn2_6/podata/1/ina/pcrUS_COMB.seq:*
6: /cgn2_6/podata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	36.2	11.5	3736	4 US-09-600-776-10
2	33	10.5	4403765	4 US-09-103-840A-2
3	33	10.5	4411529	4 US-09-107-840A-1
4	31	9.8	321	4 US-09-146-930-3
5	30.8	9.8	5159	3 US-08-458-240-3
6	30.8	9.8	5159	3 PCT-US93-03993-3
7	30.8	9.8	5159	3 PCT-US93-03993-3
8	30.2	9.6	3472	6 5244792-1
9	29.6	9.4	289	4 US-09-007-005-17
10	29.6	9.4	289	4 US-09-244-796-17
11	29.6	9.4	1824	2 US-08-822-701-1
12	29.6	9.4	1824	3 US-08-935-855-1
13	29.6	9.4	2081	3 US-08-935-855-1
14	29.6	9.4	4074	4 US-08-894-997-49
15	29.6	9.4	15202	3 US-08-922-635-21
16	29.2	9.3	7218	1 US-08-232-463-14
17	28.4	9.0	537	4 US-09-280-116-96
18	28.4	9.0	4656	4 US-09-150-460B-4
19	28.4	9.0	6614	3 US-09-150-460B-3
20	28.2	9.0	237	2 US-08-611-757-106
21	28.2	9.0	237	5 PCT-US95-05980-106
22	28.2	9.0	268	1 US-08-466-033-33
23	28.2	9.0	268	1 US-08-444-733-33
24	28.2	9.0	268	2 US-08-464-134-33
25	28.2	9.0	268	2 US-08-461-361-33
26	28.2	9.0	268	2 US-08-485-010-33
27	9.0	313	1 US-08-466-033-33	

RESULT 1

US-09-600-776-10
; Sequence 10, Application US/09600776
; Patent No. 6326168
; GENERAL INFORMATION:
; APPLICANT: Yamamotochi Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: A novel potassium channel protein.
; FILE REFERENCE: Y9903-PCT
; CURRENT APPLICATION NUMBER: US/09/600,776
; CURRENT FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: JP P1998-011434
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 1-JP P1998-346198
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 3736
; TYPE: DNA
; ORGANISM: Rattus sp.
; US-09-600-776-10

Query Match 11.5%; Score 36.2; DB 4; Length 3736;
Best Local Similarity 48.8%; Pred. No. 0.049; Mismatches 98; Conservative 0; Indexes 103; Gaps 0;

```

QY 13 TCTGAGCCATGCCATGTGCACCACTTGCCCTTCTCCCTAGGCCCTACACCCCTGCC 72
Db 3132 TCTGACCCCCAACCCCTCTGAAAGCAGCTCCAGCTGGTCAGACACATTCCACTGACC 3191
Db 3192 CTGGCCCTGGGCCAACGCCCTGGGGCTTATTAACCTGGCATCAGGGAGG 3251
QY 133 CTGGACAACTCCCTGCCCTCACCGACAGACAGATTCACTGGGACCGGGG 192
Db 3252 TGGGCTCTGGCCCTGGCTGGCTGGGTGAGCTGGCTGGCTGGCTGG 3311
Db 3312 ATTCTCTGGACTTTAACAA 3332

```

RESULT 2
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.

APPLICANT: FRASER, Claire M.
 APPLICANT: VENIER, John C.
 TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 FILE REFERENCE: 24366-2007.00
 CURRENT APPLICATION NUMBER: US/09/103, 840A
 CURRENT FILING DATE: 1998-06-24
 NUMBER OF SEQ ID NOS: 2
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO: 2
 LENGTH: 4403765
 TYPE: DNA
 ORGANISM: Mycobacterium tuberculosis
 FEATURE:
 OTHER INFORMATION: CDC 1551
 OTHER INFORMATION: "n" bases at various positions throughout the sequence
 US-09-103-840A-2
 Query Match 10.5%; Score 33; DB 4; Length 4403765;
 Best Local Similarity 55.8%; Pred. No. 8.3.; Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 Qy 33 ACACCCACTGGCTCTCTCTAGGCCACACCTGGCTCTTCTTCGCCCCCTGTACT 92
 Db 1838235 ACGCGCAGGCGCCTGTTGCGCCAAACACTGACGATGTTGGCTCTGTGTC 1838294
 Qy 93 TACTGTGGCTGGAGCATGAGCGCACTCCAAAGCATGCTGGACATAGTGC 145
 Db 1838295 CAGCGCTCCGGTAGAAATACCGGGATGTCGAGAACCGGTAACAGTGTGTC 1838247

RESULT 3
 US-09-103-840A-1
 ; Sequence 1, Application US/09103840A
 ; Patent No. 6294228
 ; GENERAL INFORMATION:
 ; APPLICANT: FLEISCHMAN, Robert D.
 ; APPLICANT: WHITE, Owen R.
 ; APPLICANT: VENIER, Claire M.
 ; APPLICANT: FRAZER, John C.
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
 ; TITLE OF INVENTION: TUBERCULOSIS
 ; FILE REFERENCE: 24366-2007.00
 ; CURRENT APPLICATION NUMBER: US/09/103, 840A
 ; CURRENT FILING DATE: 1998-06-24
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO: 1
 ; LENGTH: 4411529
 ; TYPE: DNA
 ; ORGANISM: Mycobacterium tuberculosis
 ; OTHER INFORMATION: H37Rv
 ; US-09-103-840A-1

Query Match 10.5%; Score 33; DB 4; Length 4411529;
 Best Local Similarity 55.8%; Pred. No. 8.3.; Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 33 ACACCCACTGGCTCTCTCTAGGCCACACCTGGCTCTTCTTCGCCCCCTGTACT 92
 Db 1847358 ACGCGCAGGCGCCTGTTGCGCCAAACACTGACGATGTTGGCTCTGTGTC 1847417
 Qy 93 TACTGTGGCTGGAGCATGAGCGCACTCCAAAGCATGCTGGACATAGTGC 145
 Db 1847418 CAGCGCTCCGGTAGAAATACCGGGATGTCGAGAACCGGCTAACAGTGTGTC 1847470

RESULT 4
 US-09-060-756-295
 ; Sequence 295, Application US/09060756
 ; Patent No. 6189597
 ; GENERAL INFORMATION:

Query Match 9.8%; Score 31; DB 4; Length 321;
 Best Local Similarity 55.0%; Pred. No. 0.85; Matches 61; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 Matches 61; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 Qy 4 CTGGCTTATTCGACCCATGCCATGTGCACACCCTGGCTCTCTCTAGGCTA 63
 Db 120 CGGTGTCATCCACAGCGTGGCCATGGCACACCAACTACGCCATTCTCCGACTGGGA 179
 Qy 64 CACCGTGCCTTCTTGCCCTGTAGTGTACTGTGGCTGGAGGTGATGAGC 114
 Db 180 CACCTTACCGCAGGCCACTCGAGGACTGTGTCCCACAGGGC 230

RESULT 5
 US-08-146-930-3
 ; Sequence 3, Application US/08146930
 ; Patent No. 5958764
 ; GENERAL INFORMATION:
 ; APPLICANT: Koop, Dennis R.
 ; APPLICANT: Rothnagel, Joseph A.
 ; APPLICANT: Greenlaugh, David A.
 ; TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LYON & LYON
 ; STREET: 611 West Sixth Street
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: U.S.A.
 ; ZIP: 90017
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" diskette, 1.44 Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
 ; SOFTWARE: Wordperfect (Version 5.1)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/146, 930
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA: including application
 ; PRIORITY APPLICATION NUMBER: described below: two
 ; APPLICATION NUMBER: 07/1876, 286
 ; FILING DATE: April 30, 1992
 ; APPLICATION NUMBER: No. 5958764 yet assigned (204/132)
 ; FILING DATE: October 29, 1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wartburg, Richard J.
 ; REGISTRATION NUMBER: 32,327
 ; REFERENCE/DOCKET NUMBER: 204/152
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5159 bases

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLogy: linear

MOLECULE TYPE: DNA (genomic)

US-08-146-930-3

Query Match 9.8%; Score 30.8; DB 2; Length 5159;

Best Local Similarity 50.7%; Pred. No. 3;

Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 127 CATGGCTTGGAAACATGTCCTGCCTCCACAGCAGTGACGATTACTGGGAC 186
 Db 408 CATCTCCAGGAACCCATGTCTACCAAACACCATTAAAGTCAAACAGCCACCGCT 467

Qy 187 TCGGGCAAAAGGAGACTGTGTCACAAGAACAAAGGCCAGGGCACACATGGTAC 246
 Db 468 ACAGTGCAGCTGAGCTGTCAGGCTGCTGGCTCAACCCCTCTGCTCAGCAGTGTGCG 527

Qy 247 CTGGCCCTGCCCACCCATGGCTAC 272
 Db 528 TGTCGCCCTCCAGGGCAGGGTGGC 553

RESULT 6
 US-08-458-240-3

Sequence 3, Application US/08458240

Patent No. 6143727

GENERAL INFORMATION:

APPLICANT: Roop, Dennis R.

APPLICANT: Rothnagel, Joseph A.

APPLICANT: Greenhalgh, David A.

TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS

TITLE OF INVENTION: AND METHODS OF USE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: LYON & LYON

STREET: 611 West Sixth Street

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90017

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: IBM MS-DOS (version 5.0)

SOFTWARE: WordPerfect (version 5.1)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08458, 240

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/146, 930

FILING DATE:

APPLICATION NUMBER: 07/876, 286

FILING DATE: April 30, 1992

APPLICATION NUMBER: NO. 6143727 yet assigned (204/132)

FILING DATE: October 29, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 204/152

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5159 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLogy: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

PCT-US93-03993-3

Query Match 9.8%; Score 30.8; DB 5; Length 5159;

Best Local Similarity 50.7%; Pred. No. 3;

Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Matches 74; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Qy 127 CATGGCTTGGAAACATGTCCTGCCTCCACAGCAGTGACGATTACTGGGAC 186
 Db 408 CATCTCCAGGAACCCATGTCTACCAAACACCATTAAAGTCAAACAGCCACCGCT 467

APPLICANT: Basillico, Claudio
TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE NUMBER OF SEQUENCES: 18
TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
STREET: 411 Hackensack Ave, Continental Plaza, 4th Floor
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,701
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/JAGENT INFORMATION:
NAME: Jackson, Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-002 N
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEX: 201-343-1684
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1824 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
8-822-701-1
8-935-855-1/C
SEQUENCE 1: Application US/08935855
PATENT NO. 6065485
GENERAL INFORMATION:
APPLICANT: Guttridge, Mark
APPLICANT: Basillico, Claudio
TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE NUMBER OF SEQUENCES: 22
TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

LENGTH: 2081 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO

ORIGINAL SOURCE:
 ORGANISM: Mus musculus
 US-08-935-855-21

Query Match

Best Local Similarity 9.4%; Score 29.6; DB 3; Length 2081;

Matches 56; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 58 GGCCTACACCCGTGCCCTTCTTGCCCTTGAGTACTGGCTGGATGACCGCC 117

Db 1658 GGTGACCACTAGTCCTCTGGCCCTTTTGCACTGTGCCGGTCTCAGCCCC 1599

Qy 118 ACTCCCAGCATGGCTGAAACAATCTCCCTCTGGCCCT 157

Db 1598 CTCCGGAGACAGTGCCTCTCCAGTTCCAGTGCCTC 1559

RESULT 14

US-08-894-997-49/C

; Sequence 49, Application US/08894997A

Patent No. 6270990

GENERAL INFORMATION:

APPLICANT: Schoenherr, Christopher J

TITLE OF INVENTION: NEURON RESTRICTIVE SILENCER FACTOR

FILE REFERENCE: 17810-502 NRSF

CURRENT APPLICATION NUMBER: US/08/894, 997A

CURRENT FILING DATE: 1998-01-06

EARLIER APPLICATION NUMBER: PCT/US95/02817

EARLIER FILING DATE: 1995-03-01

EARLIER APPLICATION NUMBER: 08/7398, 590

EARLIER FILING DATE: 1995-03-03

NUMBER OF SEQ ID NOS: 55

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO: 49

LENGTH: 4057

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE: gene

NAME/KEY: gene

LOCATION: (1)..(4057)

OTHER INFORMATION: Human NSRF

US-08-894-997-49

Query Match 9.4%; Score 29.6; DB 4; Length 4057;

Best Local Similarity 64.7%; Pred. No. 6.6; Matches 44; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Db 212 AGACAAAGAGCCAGGGCACATGGTACCTGCCCTCACCCATGGCA 271

Db 2818 AAGCCAGAGGAGCCAGTGTGACACTTGCTGCTGATGCTCCACCTCTGGCATGGCTGA 2759

Qy 272 CATGGTT 279

Db 2758 GTTGTGTT 2751

RESULT 15

US-08-922-635-21

; Sequence 21, Application US/08922635A

Patent No. 6033871

GENERAL INFORMATION:

APPLICANT: PILETZ, John E.

APPLICANT: IVANOV, Tima R.

TITLE OF INVENTION: DNA MOLECULES ENCODING IMPIDALINE RECEPTIVE POLYPEPTIDES

FILE REFERENCE: Corrected Sequence Listing

Patent No. 6033871
 CURRENT APPLICATION NUMBER: US/08/922, 635A
 CURRENT FILING DATE: 1997-09-03
 EARLIER APPLICATION NUMBER: 08/650, 766
 EARLIER FILING DATE: 1996-05-20
 EARLIER APPLICATION NUMBER: 60/012, 600
 EARLIER FILING DATE: 1996-03-01
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: Patentin Ver. 2.0

SEQ ID NO: 21
 LENGTH: 15202
 TYPE: DNA

ORGANISM: Homo sapiens
 US-08-922-635-21

Query Match 9.4%; Score 29.6; DB 3; Length 15202;

Matches 59; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 16 GAGCCATGGCATGGCACACCCACTGGCCCTCTCTCTAGGCTTACACCCGGCCCT 75

Db 8372 GGCCCTCCAGCATGGCCACTGTGCTGGCTCTCTGACACGCCAGACATGCCGG 8479

Qy 76 TCTTGGCCCTTGATCTACTGTGCTGGAGTGATGAGGCCACTGCC 123

Db 8432 CGGGAGGTCAGGTGGCCAGATCATCTCCCTGCC 8479

Search completed: June 21, 2003, 03:56:25
 Job time : 46 secs

Copyright (c) 1993 - 2003 Compugen Ltd.
GenCore version 5.1.6

OM nucleic - nucleic search, using sw model
Run on: June 21, 2003, 03:25:03 ; Search time 157 Seconds
(without alignments)
4518.341 Million cell updates/sec

Title: US-09-955-807-1_COPY_59_373
Perfect score: 315
Sequence: 1 atgcttgggttattcttgagcc.....atgtgttttcagcgagcc 315
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:

4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Result No.	Score	Query Match Length	DB ID	Description
1	315	100.0	1692	Human secretory pr
2	313.4	99.5	1201	Human secretory pr
3	304	96.5	384	Human secretory pr
4	39.6	12.6	521	Murine 7-transmemb
5	36.2	11.5	3736	Rat brain specific
6	36.2	11.5	10732	Gene encoding a su
7	33	10.5	9545	Human immune/haema
8	33	10.5	4403765	Mycobacterium tube
9	33	10.5	4411529	Mycobacterium tube

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

			Description
1	315	100.0	Human secretory pr
2	313.4	99.5	Human secretory pr
3	304	96.5	Human secretory pr
C	4	39.6	Murine 7-transmemb
C	5	36.2	Rat brain specific
C	6	36.2	Gene encoding a su
C	7	33	Hum. immune/haema
C	8	33	Mycobacterium tube
C	9	33	Mycobacterium tube
	10.5	4411529	AAT199682

ALIGNMENT:

DNA encoding human	
Human nervous syst	
CERP-B cDNA.	Homo
DNA encoding centr	
DNA encoding human	
Human DNA sequence	
DNA encoding novel	
Human calpain CAPN	
cDNA sequence #204	
Human ORF1 PolyNUC	
Human T-type volta	
Human T-type volta	
Human poly(adenylate)	
Drosophila melanog	
Drosophila melanog	
Drosophila melanog	
Human prostate exp	
Human 7-transmemb	
Drosophila melanog	
Human BBSP RPR CD	
Human nuclear rece	
Drosophila melanog	
Human immune haeme	
Human poly(adenylate)	
Drosophila melanog	
Drosophila odorant	
Drosophila melanog	
Drosophila melanog	
Mycobacterium tube	
Human NF- κ B activa	
Novel human coding	
Human full-length	

XX
XX 01-OCT-1998; 98US-0164740.

PR ID AAZ4892 standard; DNA; 12001 BP.

XX XX

PA AC

XX KW

PI DT 01-AUG-2000 (first entry)

XX DE Human secretory protein 48 (Zsig48) genomic DNA.

XX KW Secretory protein 48; Zsig48; human; leukocyte; proliferation;

PT KW immunostimulant; adjuvant; immunomodulator; gene therapy;

PT PT individuals -

XX PS Example 1; Page 127-128; 141pp; English.

XX CC The present sequence is that of cDNA coding for human secretory

protein 48 (Zsig48, see AAY9436), a protein that stimulates the

proliferation of peripheral blood mononuclear cells, i.e. T-cells,

B-cells and monocytes. An expressed sequence tag (see AAK94890) was

isolated through the random sequencing of a mixed haematopoietic

cDNA library, and the present full-length cDNA was subsequently

obtained. The Zsig48 gene was mapped to human chromosome 7q36.3.

The invention provides Zsig48 polynucleotides, polypeptides,

antibodies and anti-idiotypic antibodies. Zsig48 stimulates the

proliferation of leukocytes in both a mixed leukocyte reaction and

in an unmixed leukocyte reaction. It can be used to promote

proliferation of peripheral blood leukocytes and to treat low

leukocyte counts in individuals (claimed). This can be useful in

treating cancer patients whose leukocytes have been depleted by

chemotherapy, radiation or illness. Zsig48 can be administered to

patients receiving bone marrow transplants to promote proliferation

of leukocytes produced by the transplanted marrow. It could also

be useful in treating immunosuppressed individuals such as the

elderly or HIV infected individuals, or used as a vaccine adjuvant.

Zsig48 polynucleotides can be used in the recombinant production of

Zsig48 polypeptides, as probes and primers to detect Zsig48 gene

expression or to identify mutations in the Zsig48 gene, in

diagnostic assays, or introduced into a subject to enhance

immunological responses (immunomodulator gene therapy).

CC Sequence 1692 BP; 490 A; 393 C; 362 G; 447 T; 0 other:

Query Match 100.0%; Score 315; DB 21; Length 1692;
Best Local Similarity 100.0%; Pred. No. 2.1e-88; Mismatches 315; Conservative 0; Indels 0; Gaps 0;

CC Qy 1 ATGCCTGGTATTCTGAGGCCATGCCATGGCACACCACTGGCCCTCTCCCTTAGGC 60

Db 59 ATGCCTGGTATTCTGAGGCCATGCCATGGCACACCACTGGCCCTCTCCCTAGGC 118

Db 61 CTACACCCCTGCCCTTGACTGTGGAGTGAGCTGGAGCTGGAGCTGGAGTGAGCCACT 120

Db 119 CTACACCCCTGCCCTTGACTGTGGAGCTGGAGTGAGCCACT 178

Qy 121 CCCAAGCATGGCCCTGAAACATGTCCTCGCCCTCCACCAAGCAGTGGAGATTCACT 180

Db 179 CCCAAGCATGGCCCTGAAACATGTCCTCGCCCTCCACCAAGCAGTGGAGATTCACT 238

Qy 181 GGAACTCGGGCAAAGGAGACTGTGTCACAGACAAGAACAGGCCAGGGTACACATG 240

Db 239 GGAACTCGGGCAAAGGAGACTGTGTCACAGACAAGAACAGGCCAGGGTACACATG 298

Qy 241 TGTACCCCTGCCCTACCCATGGCTGACATGGTGACACCTCTGAAATCATGTG 300

Db 299 TGTACCCCTGCCCTACCCATGGCTGACATGGTGACACCTCTGAAATCATGTG 358

Qy 301 TCTTCAGGCCAGCCAC 315

Db 359 TCTTCAGGCCAGCCAC 373

AAZ4892 ID AAZ4892 standard; DNA; 12001 BP.

XX XX

PA AC

XX KW

PI DT 01-AUG-2000 (first entry)

XX DE Human secretory protein 48 (Zsig48) genomic DNA.

XX KW Secretory protein 48; Zsig48; human; leukocyte; proliferation;

PT KW immunostimulant; adjuvant; immunomodulator; gene therapy;

PT PT individuals -

XX PS Example 1; Page 127-128; 141pp; English.

XX CC The present sequence is that of genomic DNA coding for human

secretory protein 48 (Zsig48, see AAY9436), a protein that stimulates

the proliferation of peripheral blood mononuclear cells, i.e. T-cells,

B-cells and monocytes. The gene maps to human chromosome 7q36.3.

CC The invention provides Zsig48 polynucleotides, polypeptides,

antibodies and anti-idiotypic antibodies. Zsig48 stimulates the

proliferation of leukocytes in both a mixed leukocyte reaction and

in an unmixed leukocyte reaction. It can be used to promote

proliferation of peripheral blood leukocytes and to treat low

leukocyte counts in individuals (claimed). This can be useful in

treating cancer patients whose leukocytes have been depleted by

chemotherapy, radiation or illness. Zsig48 can be administered to

patients receiving bone marrow transplants to promote proliferation

of leukocytes produced by the transplanted marrow. It could also

be useful in treating immunosuppressed individuals such as the

elderly or HIV infected individuals, or used as a vaccine adjuvant.

Zsig48 polynucleotides can be used in the recombinant production of

Zsig48 polypeptides, as probes and primers to detect Zsig48 gene

expression or to identify mutations in the Zsig48 gene, in

diagnostic assays, or introduced into a subject to enhance

immunological responses (immunomodulator gene therapy).

CC Sequence 12001 BP; 3336 A; 2441 C; 2461 G; 3763 T; 0 other:

Query Match 99.5%; Score 313.4; DB 21; Length 12001;
Best Local Similarity 99.7%; Pred. No. 1.3e-87; Mismatches 314; Conservative 0; Indels 1; Gaps 0;

CC Qy 1 ATGCCTGGTATTCTGAGGCCATGCCATGGCACACCCACTGGCCCTCTCCCTAGGC 60

us-09-955-807-1_copy_59_373.rng

RESULT 7	
AAK8011	PR
ID	PR
AAK8011 standard; DNA; 9545 BP.	PR
XX	PR
AC	PR
AAK8011;	PR
XX	PR
DT	PR
07-NOV-2001 (first entry)	PR
Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34923.	PR
Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; cytostatic; gene therapy; vaccine; metastasis; ds.	PR
KW	PR
XX	PR
OS	PR
Homo sapiens.	PR
XX	PR
WO200157182-A2.	PR
XX	PR
PD	PR
09-AUG-2001.	PR
XX	PR
PF	PR
17-JAN-2001; 2001WO-US01354.	PR
XX	PR
PR	PR
31-JAN-2000; 2000US-0179065.	PR
PR	PR
04-FEB-2000; 2000US-0180628.	PR
PR	PR
24-FEB-2000; 2000US-0186664.	PR
PR	PR
02-MAR-2000; 2000US-0186350.	PR
PR	PR
16-MAR-2000; 2000US-0189874.	PR
PR	PR
17-MAR-2000; 2000US-019076.	PR
PR	PR
18-APR-2000; 2000US-0198123.	PR
PR	PR
19-MAY-2000; 2000US-0205515.	PR
PR	PR
07-JUN-2000; 2000US-0204667.	PR
PR	PR
28-JUN-2000; 2000US-0214886.	PR
PR	PR
30-JUN-2000; 2000US-0215335.	PR
PR	PR
07-JUL-2000; 2000US-0216647.	PR
PR	PR
07-JUL-2000; 2000US-0216880.	PR
PR	PR
11-JUL-2000; 2000US-0214887.	PR
PR	PR
11-JUL-2000; 2000US-0217496.	PR
PR	PR
14-JUL-2000; 2000US-0218290.	PR
PR	PR
26-JUL-2000; 2000US-0220963.	PR
PR	PR
26-JUL-2000; 2000US-0220964.	PR
PR	PR
14-AUG-2000; 2000US-0224518.	PR
PR	PR
14-AUG-2000; 2000US-0224519.	PR
PR	PR
14-AUG-2000; 2000US-022513.	PR
PR	PR
14-AUG-2000; 2000US-025114.	PR
PR	PR
14-AUG-2000; 2000US-022566.	PR
PR	PR
14-AUG-2000; 2000US-022567.	PR
PR	PR
14-AUG-2000; 2000US-025668.	PR
PR	PR
14-AUG-2000; 2000US-025670.	PR
PR	PR
14-AUG-2000; 2000US-0225447.	PR
PR	PR
14-AUG-2000; 2000US-022557.	PR
PR	PR
14-AUG-2000; 2000US-022558.	PR
PR	PR
14-AUG-2000; 2000US-022559.	PR
PR	PR
18-AUG-2000; 2000US-026379.	PR
PR	PR
22-AUG-2000; 2000US-0226681.	PR
PR	PR
12-SEP-2000; 2000US-0231168.	PR
PR	PR
14-SEP-2000; 2000US-0232397.	PR
PR	PR
14-SEP-2000; 2000US-0232398.	PR
PR	PR
14-SEP-2000; 2000US-0232399.	PR
PR	PR
14-SEP-2000; 2000US-0232401.	PR
PR	PR
14-SEP-2000; 2000US-0233063.	PR
PR	PR
14-SEP-2000; 2000US-0233065.	PR
PR	PR
21-SEP-2000; 2000US-0234223.	PR
PR	PR
21-SEP-2000; 2000US-0234274.	PR
PR	PR
25-SEP-2000; 2000US-0234597.	PR
PR	PR
25-SEP-2000; 2000US-0235064.	PR
PR	PR
26-SEP-2000; 2000US-0235984.	PR
PR	PR
27-SEP-2000; 2000US-0235934.	PR
PR	PR
29-SEP-2000; 2000US-0235836.	PR
PR	PR
29-SEP-2000; 2000US-0236367.	PR
PR	PR
29-SEP-2000; 2000US-0236368.	PR
PR	PR
29-SEP-2000; 2000US-0236359.	PR
PR	PR
02-OCT-2000; 2000US-0236802.	PR
PR	PR
02-OCT-2000; 2000US-0237038.	PR
PR	PR
02-OCT-2000; 2000US-0237039.	PR
PR	PR
02-OCT-2000; 2000US-0237040.	PR
PR	PR
13-OCT-2000; 2000US-0239925.	PR
PR	PR
20-OCT-2000; 2000US-0239937.	PR
PR	PR
20-OCT-2000; 2000US-0240560.	PR
PR	PR
20-OCT-2000; 2000US-0241221.	PR
PR	PR
20-OCT-2000; 2000US-0241785.	PR
PR	PR
20-OCT-2000; 2000US-0237040.	PR
PR	PR
20-OCT-2000; 2000US-0239925.	PR
PR	PR
20-OCT-2000; 2000US-0239935.	PR
PR	PR
20-OCT-2000; 2000US-0239937.	PR
PR	PR
20-OCT-2000; 2000US-0241809.	PR
PR	PR
20-OCT-2000; 2000US-0241826.	PR
PR	PR
01-NOV-2000; 2000US-0244617.	PR
PR	PR
08-NOV-2000; 2000US-0246774.	PR
PR	PR
08-NOV-2000; 2000US-0246776.	PR
PR	PR
08-NOV-2000; 2000US-0246776.	PR
PR	PR
08-NOV-2000; 2000US-0246477.	PR
PR	PR
08-NOV-2000; 2000US-0246477.	PR
PR	PR
08-NOV-2000; 2000US-0246478.	PR
PR	PR
08-NOV-2000; 2000US-0246523.	PR
PR	PR
08-NOV-2000; 2000US-0246525.	PR
PR	PR
08-NOV-2000; 2000US-0246526.	PR
PR	PR
08-NOV-2000; 2000US-0246527.	PR
PR	PR
08-NOV-2000; 2000US-0246528.	PR
PR	PR
08-NOV-2000; 2000US-0246532.	PR
PR	PR
08-NOV-2000; 2000US-0246532.	PR
PR	PR
08-NOV-2000; 2000US-0246610.	PR
PR	PR
08-NOV-2000; 2000US-0246611.	PR
PR	PR
08-NOV-2000; 2000US-0246613.	PR
PR	PR
17-NOV-2000; 2000US-0246927.	PR

RESULT 8
 PR 17-NOV-2000; 200005-0249209.
 PR 17-NOV-2000; 200005-0249210.
 PR 17-NOV-2000; 200005-0249211.
 PR 17-NOV-2000; 200005-0249212.
 PR 17-NOV-2000; 200005-0249213.
 PR 17-NOV-2000; 200005-0249214.
 PR 17-NOV-2000; 200005-0249215.
 PR 17-NOV-2000; 200005-0249216.
 PR 17-NOV-2000; 200005-0249217.
 PR 17-NOV-2000; 200005-0249218.
 PR 17-NOV-2000; 200005-0249219.
 PR 17-NOV-2000; 200005-0249220.
 PR 17-NOV-2000; 200005-0249224.
 PR 17-NOV-2000; 200005-0249245.
 PR 17-NOV-2000; 200005-0249264.
 PR 17-NOV-2000; 200005-0249265.
 PR 17-NOV-2000; 200005-0249297.
 PR 17-NOV-2000; 200005-0249299.
 PR 01-DEC-2000; 200005-0249300.
 PR 01-DEC-2000; 200005-0249310.
 PR 01-DEC-2000; 200005-0249319.
 PR 05-DEC-2000; 200005-0251030.
 PR 05-DEC-2000; 200005-0251988.
 PR 05-DEC-2000; 200005-0251979.
 PR 01-DEC-2000; 200005-0251856.
 PR 08-DEC-2000; 200005-0251868.
 PR 08-DEC-2000; 200005-0251869.
 PR 01-DEC-2000; 200005-0251989.
 PR 08-DEC-2000; 200005-0251990.
 PR 11-DEC-2000; 200005-0254097.
 PR 05-JAN-2001; 200005-0259578.

PA (HUMA-) HUMAN GENOME SCI. INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI; 2001-483426/52.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis - Disclosure; SEQ ID NO 34923; 3071pp + Sequence Listing; English.
 PT AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic activities, and may be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK6703 to AAK67694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK4942 to AAK54950 and AAM82169 represent sequences used in the exemplification of the present invention.

SQ Sequence 9545 BP; 2568 A; 2101 C; 2253 G; 2623 T; 0 other;

Query Match 33 ACACCCACTGGCCTTTCCTTAGGCCACCTGCCCCTTCTTGCCCTGTAG Score 10.5%; Best Local Similarity 55.8%; Length 4403765; Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 1838295 CAGCGTTCGGCTAGGATCACGGGATGAGCACACCGCTAAACAGTC ID AA199683; Name AA199683 standard; DNA; 4403765 BP; DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2. XX KW Mycobacterium tuberculosis; strain H37Rv; genome; variation; epidemiology; patient treatment; epidemic monitoring; ds. XX OS Mycobacterium tuberculosis.

PR 15-JAN-2002 (first entry)
 XX PN US6294328-B1.
 XX PR 24-JUN-1998; 98US-0103840.
 XX PA (GENO-) INST GENOMIC RES.
 XX PT Fleischmann RD, White OR, Fraser CM, Venter JC; DR WPI; 2001-647261/74.
 XX PT Evaluating strain variation of *Mycobacterium tuberculosis*, comprises determining the nucleotide sequence of the strain at positions in the genome corresponding to positions where *M. tuberculosis* strains CDC PR 1551 and H37Rv differ.
 XX PS Claim 4; SEQ ID NO 2; 3PP + Sequence Listing; English.
 XX CC The invention relates to evaluating strain variation within and between different populations of the tuberculosis bacterial pathogen, *Mycobacterium tuberculosis* or related *Mycobacterium* by determining the nucleotide sequence of the first strain at positions in the complete sequence of the genome that correspond to positions that differ in the nucleotide sequences of *M. tuberculosis* strains CDC 1551 (AA199683) and H37Rv (AA199682). The method is useful for evaluating strain variation of *M. tuberculosis* and has valuable application in the fields of tuberculosis genetics, epidemiology, patient treatment and epidemic monitoring.
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docid=62943281.
 XX SQ Sequence 4403765 BP; 757105 A; 1447799 C; 141301 G; 757371 T; 189 other;

Query Match 93 TACTGGCTGAGCTGATGCCGCACTCCAGATGGCTTGGACATGTC Score 9.9%; Best Local Similarity 55.8%; Length 145; Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 1838295 CAGCGTTCGGCTAGGATCACGGGATGAGCACACCGCTAAACAGTC ID AA199683; Name AA199683 standard; DNA; 4403765 BP; DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2. XX KW Mycobacterium tuberculosis; strain H37Rv; genome; variation; epidemiology; patient treatment; epidemic monitoring; ds. XX OS Mycobacterium tuberculosis.

PR 25-SEP-2001.
 XX PP 24-JUN-1998; 98US-0103840.
 XX PR 24-JUN-1998; 98US-0103840.
 XX PA (GENO-) INST GENOMIC RES.
 XX PT Fleischmann RD, White OR, Fraser CM, Venter JC; DR WPI; 2001-647261/74.
 XX PT Evaluating strain variation of *Mycobacterium tuberculosis*, comprises determining the nucleotide sequence of the strain at positions in the genome corresponding to positions where *M. tuberculosis* strains CDC PR 1551 and H37Rv differ.
 XX PS Claim 4; SEQ ID NO 2; 3PP + Sequence Listing; English.
 XX CC The invention relates to evaluating strain variation within and between different populations of the tuberculosis bacterial pathogen, *Mycobacterium tuberculosis* or related *Mycobacterium* by determining the nucleotide sequence of the first strain at positions in the complete sequence of the genome that correspond to positions that differ in the nucleotide sequences of *M. tuberculosis* strains CDC 1551 (AA199683) and H37Rv (AA199682). The method is useful for evaluating strain variation of *M. tuberculosis* and has valuable application in the fields of tuberculosis genetics, epidemiology, patient treatment and epidemic monitoring.
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?docid=62943281.
 XX SQ Sequence 4403765 BP; 757105 A; 1447799 C; 141301 G; 757371 T; 189 other;

Query Match 93 TACTGGCTGAGCTGATGCCGCACTCCAGATGGCTTGGACATGTC Score 9.9%; Best Local Similarity 55.8%; Length 145; Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

OY 1838295 CAGCGTTCGGCTAGGATCACGGGATGAGCACACCGCTAAACAGTC ID AA199683; Name AA199683 standard; DNA; 4403765 BP; DE Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 2. XX KW Mycobacterium tuberculosis; strain H37Rv; genome; variation; epidemiology; patient treatment; epidemic monitoring; ds. XX OS Mycobacterium tuberculosis.

DB Mycobacterium tuberculosis strain H37Rv genome SEQ ID NO 1.
 XX
 KW Mycobacterium tuberculosis; strain H37Rv; genome CPC 1551; genome variation; epidemiology; patient treatment; epidemic monitoring; ds.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN US6294328-B1.
 XX
 PD 25-SEP-2001.
 XX
 PR 24-JUN-1998; 98US-0103840.
 XX
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Fleischmann RD, White OR, Fraser CM, Venter JC;
 XX
 DR WPI; 2001-647261/74.
 XX
 PT Evaluating strain variation of Mycobacterium tuberculosis, comprises determining the nucleotide sequence of the strain at positions in the genome corresponding to positions where M. tuberculosis strains CPC PT 1551 and H37Rv differ -
 XX
 PS Claim 3; SEQ ID NO 1; 3pp + sequence Listing; English.
 CC The invention relates to evaluating strain variation within and between different populations of the tuberculosis bacteria; pathogen, Mycobacterium tuberculosis or related Mycobacterium by determining the nucleotide sequence of the first strain at positions in the complete sequence of the genome that correspond to positions that differ in the nucleotide sequences of M. tuberculosis strains CPC 1551 (ATCC9968) and H37Rv (ATCC9682). The method is useful for evaluating strain variation of M. tuberculosis and has valuable application in the fields of tuberculosis genetics, epidemiology, patient treatment and epidemic monitoring.
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPRO at seqdata.uspto.gov/sequence.html?docid=6294328B1.
 XX
 SQ Sequence 4411529 BP; 75855 A; 1449883 C; 1444602 G; 758379 T; 0 other;
 XX
 Query Match 10.5%; Score 33; DB 22; Length 4411529;
 Best Local Similarity 55.8%; Pred. No. 54;
 Matches 63; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
 QY 33 ACACCCACTTGACCTCTTGTAGGCTTACACCCGCCATTCTGGCTTGAGT 92
 Db 1847358 ACCGCAGACGGCTTGTTGCGCAGAACCTGACGATGGTTGCGCTTGTC 1847417
 QY 93 TACTGTGCTGGAGTGTAGGCCCCACTCCCAANGCATGGCAATGTCCTCGC 152
 Db 203 CCCCTGCAGGGCTTCGGGCCACCTGGCGCACGGCACGGGACACTCCGGCTGC 262
 QY 153 CCCCTGCACCGAGT 167
 Db 263 GCTGCCCTGGCACT 277

RESULT 10
 ABBK51611
 ID ABBK51611 standard; DNA; 378 BP.
 XX
 AC ABBK51611;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE DNA encoding human sulfotransferase, fragment #3.
 XX
 KW Sulfotransferase; neuroprotective; sulfotransferase modulator; enzyme therapy; cancer; central nervous system disorder; CNS disorder; diabetes; asthma; chronic obstructive pulmonary disease; COPD; human; ds.
 XX
 OS Homo sapiens.

RESULT 11
 ABB20784/C
 ID ABB20784 standard; DNA; 2633 ·BP.
 XX
 AC ABB20784;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 13115.
 XX
 KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary; antiparkinsonian; antisickling; antiaemic; antiarthritic; cancer; antiheumatic; hepatoprotective; cerebroprotective; antiinflammatory; antiallergic; antidiabetic; antilulcer; anticonvulsant; antifungal; antiparasitic; cardiotonic; immune disorder; cardiovascular disorder;

KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds. PR 25-SEP-2000; 2000US-0234998.
 XX OS Homo sapiens. PR 26-SEP-2000; 2000US-0234984.
 XX PN WO200159063-A2. PR 27-SEP-2000; 2000US-0238834.
 XX PD 16-AUG-2001. PR 29-SEP-2000; 2000US-0230367.
 XX 17-JAN-2001; 2001WO-US01334. PR 29-SEP-2000; 2000US-0230368.
 XX PR 29-SEP-2000; 2000US-0230369.
 PR 31-JAN-2000; 2000US-0179065. PR 29-SEP-2000; 2000US-0230802.
 PR 04-FEB-2000; 2000US-0180628. PR 02-OCT-2000; 2000US-0230307.
 PR 24-FEB-2000; 2000US-0184664. PR 02-OCT-2000; 2000US-0230040.
 PR 02-MAR-2000; 2000US-0186350. PR 13-OCT-2000; 2000US-0230935.
 PR 16-MAR-2000; 2000US-0188874. PR 13-OCT-2000; 2000US-0230937.
 PR 17-MAR-2000; 2000US-0190076. PR 20-OCT-2000; 2000US-0230960.
 PR 18-APR-2000; 2000US-0198123. PR 20-OCT-2000; 2000US-0240960.
 PR 19-MAY-2000; 2000US-0205515. PR 20-OCT-2000; 2000US-0241785.
 PR 07-JUN-2000; 2000US-0203467. PR 20-OCT-2000; 2000US-0241786.
 PR 28-JUN-2000; 2000US-0214886. PR 20-OCT-2000; 2000US-0241787.
 PR 30-JUN-2000; 2000US-021515. PR 20-OCT-2000; 2000US-0241808.
 PR 07-JUL-2000; 2000US-0216647. PR 20-OCT-2000; 2000US-0241809.
 PR 07-JUL-2000; 2000US-0216880. PR 20-OCT-2000; 2000US-0241826.
 PR 11-JUL-2000; 2000US-0217487. PR 20-OCT-2000; 2000US-0242221.
 PR 11-JUL-2000; 2000US-0217496. PR 01-NOV-2000; 2000US-0241617.
 PR 14-JUL-2000; 2000US-0221213. PR 08-NOV-2000; 2000US-0244524.
 PR 26-JUL-2000; 2000US-0225266. PR 08-NOV-2000; 2000US-0244745.
 PR 26-JUL-2000; 2000US-0226964. PR 08-NOV-2000; 2000US-024476.
 PR 14-AUG-2000; 2000US-0224518. PR 08-NOV-2000; 2000US-024477.
 PR 14-AUG-2000; 2000US-0224519. PR 08-NOV-2000; 2000US-024478.
 PR 14-AUG-2000; 2000US-0225213. PR 08-NOV-2000; 2000US-0245523.
 PR 14-AUG-2000; 2000US-0225214. PR 08-NOV-2000; 2000US-0245524.
 PR 14-AUG-2000; 2000US-0225758. PR 08-NOV-2000; 2000US-0245525.
 PR 14-AUG-2000; 2000US-0225759. PR 08-NOV-2000; 2000US-0245526.
 PR 18-AUG-2000; 2000US-0226267. PR 08-NOV-2000; 2000US-0245527.
 PR 22-AUG-2000; 2000US-0225270. PR 08-NOV-2000; 2000US-0245528.
 PR 14-AUG-2000; 2000US-0225447. PR 08-NOV-2000; 2000US-0245609.
 PR 14-AUG-2000; 2000US-0225757. PR 08-NOV-2000; 2000US-0246110.
 PR 14-AUG-2000; 2000US-0225758. PR 08-NOV-2000; 2000US-0246111.
 PR 14-AUG-2000; 2000US-0226267. PR 08-NOV-2000; 2000US-024613.
 PR 14-AUG-2000; 2000US-0226279. PR 08-NOV-2000; 2000US-0249207.
 PR 22-AUG-2000; 2000US-0226688. PR 17-NOV-2000; 2000US-0249208.
 PR 22-AUG-2000; 2000US-0227182. PR 17-NOV-2000; 2000US-0249209.
 PR 23-AUG-2000; 2000US-0227009. PR 17-NOV-2000; 2000US-0249210.
 PR 30-AUG-2000; 2000US-0228759. PR 17-NOV-2000; 2000US-0249211.
 PR 01-SEP-2000; 2000US-0228924. PR 17-NOV-2000; 2000US-0249212.
 PR 01-SEP-2000; 2000US-0229343. PR 17-NOV-2000; 2000US-0249213.
 PR 01-SEP-2000; 2000US-0229344. PR 17-NOV-2000; 2000US-0249214.
 PR 05-SEP-2000; 2000US-0229345. PR 17-NOV-2000; 2000US-0249215.
 PR 05-SEP-2000; 2000US-0229509. PR 17-NOV-2000; 2000US-0249216.
 PR 05-SEP-2000; 2000US-0229513. PR 17-NOV-2000; 2000US-0249217.
 PR 06-SEP-2000; 2000US-0230437. PR 17-NOV-2000; 2000US-0249218.
 PR 06-SEP-2000; 2000US-0230438. PR 17-NOV-2000; 2000US-0249219.
 PR 08-SEP-2000; 2000US-0231242. PR 17-NOV-2000; 2000US-0249224.
 PR 08-SEP-2000; 2000US-0231243. PR 17-NOV-2000; 2000US-0249225.
 PR 08-SEP-2000; 2000US-0231244. PR 17-NOV-2000; 2000US-0249226.
 PR 08-SEP-2000; 2000US-0231413. PR 17-NOV-2000; 2000US-0249227.
 PR 08-SEP-2000; 2000US-0231414. PR 17-NOV-2000; 2000US-0249228.
 PR 08-SEP-2000; 2000US-0232080. PR 17-NOV-2000; 2000US-0249300.
 PR 14-SEP-2000; 2000US-0232400. PR 01-DEC-2000; 2000US-0250391.
 PR 14-SEP-2000; 2000US-0232401. PR 01-DEC-2000; 2000US-0251160.
 PR 14-SEP-2000; 2000US-0232402. PR 05-DEC-2000; 2000US-0251030.
 PR 14-SEP-2000; 2000US-0232403. PR 05-DEC-2000; 2000US-0251988.
 PR 14-SEP-2000; 2000US-0232404. PR 05-DEC-2000; 2000US-0256719.
 PR 14-SEP-2000; 2000US-0232405. PR 05-DEC-2000; 2000US-0257199.
 PR 14-SEP-2000; 2000US-0232406. PR 08-DEC-2000; 2000US-0251868.
 PR 14-SEP-2000; 2000US-0232407. PR 08-DEC-2000; 2000US-0251869.
 PR 14-SEP-2000; 2000US-0232408. PR 08-DEC-2000; 2000US-0251989.
 PR 21-SEP-2000; 2000US-0233065. PR 08-DEC-2000; 2000US-0251990.
 PR 21-SEP-2000; 2000US-0233223. PR 11-DEC-2000; 2000US-0250907.
 PR 25-SEP-2000; 2000US-0234997. PR 05-JAN-2001; 2001US-0259678.

XX
PA
XX
PI Shiffman D, Somogyi R, Lawn R, Seilhamer JJ, Porter GJ, Mikita T;
PT Tai J;
XX
DR WPI; 2002-010925/01.

XX
PT Composition useful for diagnosis of conditions, disorders or diseases
associated with atherosclerosis, comprises several polynucleotides that
are differentially expressed in foam cell development
XX
PS Claim 1; Page 130-131; 315pp; English.
XX
CC The present invention relates to the isolation of human polynucleotide
CC sequences that are differentially expressed during foam cell
CC differentiation. The polynucleotide sequences of the invention or a
CC composition comprising these polynucleotides are useful as a high
CC throughput method for detecting altered expression of one or more
CC polynucleotides in a sample. The polynucleotides can be used in the
CC diagnosis of disorders associated with foam cell development such as
CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
CC coronary artery disease. The polynucleotide sequences can also be used
CC as PCR primers and probes. The polynucleotides of the invention are also
CC useful in gene therapy. AAS94746-AAS95021 represent the human
CC polynucleotide sequences of the invention which are differentially
CC expressed during foam cell differentiation.

XX
Sequence 4748 BP; 1123 A; 1239 C; 1125 G; 1258 T; 3 other;

Query Match 10.3%; Score 32.6; DB 24; Length 4748;
Best Local Similarity 52.6%; Pred. No. 6, 9;
Matches 71; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 33 ACACCCACTTGCCCTCTTCTAGGCCCTAACCCCTGCCCTCTCTGCCCCCTGTAGT 92
DB 1216 ACCCTACTCTACCAACTCCCTACGGCTCTACTCCAAAGCGCTTCGGCTCCACCTTCGACG 1275
QY 93 TACTGTGGCTGGATGATGAGGCCCACTTCCAAAGCATGGCCTGAACTATGNCCTCGGC 152
DB 1276 CCTTGCGCAAGGCCCTCTGGGCCACCTGGCGCACGG3AAGCACTTCGGCTGC 1335
QY 153 CCCTCCACCAAGGT 167
DB 1336 GCCTGCCCTCCGCACT 1350

Search completed: June 21, 2003. 03:55:34
Job time : 163 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: June 17, 2003, 10:25:41 ; Search time 22 Seconds
Sequence: (without alignments)
Scoring table: BLOSUM62
Gapcp 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	80	13.7	264	1 SNAI_MOUSE	Q02085 mus musculus
2	77	13.2	388	1 SOX7_HUMAN	Q0bt81 homo sapiens
3	76	13.1	380	1 SOX7_MOUSE	P40646 mus musculus
4	73.5	12.6	323	1 OIG2_HUMAN	Q13516 homo sapiens
5	72	12.4	375	1 M15_PIG	P79295 sus scrofa
6	71.5	12.3	393	1 DIP4_RAT	Q9y2h0 homo sapiens
7	71	12.2	992	1 DIP4_RAT	P77839 rattus norvegicus
8	69	11.9	1964	1 NTC4_MOUSE	P31695 mus musculus
9	69	11.9	762	1 E13B_TRIMA	P53626 trichodermella
10	69	1032	1 S24D_HUMAN	Q94855 homo sapiens	
11	68.5	11.8	323	1 OIG2_MOUSE	Q9eqw6 mus musculus
12	67	11.5	357	1 MTD1_ARATH	P002971 arabidopsis thaliana
13	67	11.5	2109	1 PCCA_CHICK	P07898 gallus gallus
14	66.5	11.4	1841	1 CC12_SHRPO	Q10509 schizosaccharomyces pombe
15	66	11.3	368	1 G162_HUMAN	E13B_TRIMA
16	66	11.3	686	1 VELH_PRVN3	Q00660 pseudorabies virus
17	66	11.3	829	1 CADG_HUMAN	O75309 homo sapiens
18	66	11.3	913	1 PCK5_HUMAN	Q92824 homo sapiens
19	66	11.3	1877	1 PCK5_MOUSE	Q04592 mus musculus
20	66	11.3	1877	1 PCK5_RAT	P14143 rattus norvegicus
21	66	11.3	3866	1 H2X_MOUSE	P55200 mus musculus
22	65.5	11.3	3685	1 DMD_HUMAN	P11532 homo sapiens
23	65	11.2	686	1 VGH_PRVKA	P21416 pseudorabies virus
24	65	11.2	1324	1 IRS2_HUMAN	O79422 homo sapiens
25	65	11.2	2259	1 CCAE_RABBIT	Q02343 oryctolagus cuniculus
26	64.5	11.1	267	1 THIG_AQUAE	Q67926 aquifex aeolicus
27	64.5	11.1	380	1 DAPA_MAZIZE	P26259 zea mays (mexican corn)
28	64.5	11.1	875	1 NEPR_HUMAN	P356730 homo sapiens
29	64.5	11.1	1077	1 AT10_HUMAN	Q9h324 homo sapiens
30	64	11.0	380	1 VASP_HUMAN	P50552 homo sapiens
31	64	11.0	445	1 GUDH_ECOLI	P276637 escherichia coli
32	64	11.0	607	1 HENKA_COW1	Q86601 canis lupus familiaris
33	63.5	10.9	245	1 HZB4_CHICK	P48480 gallus gallus

ALIGNMENTS

RESULT 1	ID	SNAI_MOUSE	STANDARD;	PRT;	264 AA.
RN	[2]	RN			
RP	SEQUENCE FROM N.A.	RX			
RA	MEDLINE-9313072; PubMed-1483390;	RA			
RT	Nieto A.M., Bennett M.F., Sargent M.G., Wilkinson D.G.; "Cloning and developmental expression of sna, a murine homologue of the Drosophila snail gene.";	RT			
RL	Development 116:227-237(1992).	RL			
RN	[3]	RN			
RP	SEQUENCE OF 1-27 FROM N.A.	RP			
RT	STRAIN=129/Sv:	RT			
RA	MEDLINE-97419272; PubMed-9271672;	RA			
RT	Jiang R., Copeland N.G., Gilbert D.J., Jenkins N.A., Gridley T.; "Genomic organization and chromosomal localization of the mouse snail (Sna) gene.";	RT			
RL	Mamm. Genome 8:686-688(1997).	RL			
CC	-- FUNCTION: THIS PROTEIN HAS MANY ROLES DURING POSTIMPLANTATION DEVELOPMENT. IT IS INVOLVED IN EMBRYONIC MESODERM FORMATION AND ITS MAINTENANCE AND MAY ALSO BE INVOLVED IN CHONDROGENESIS AND IN EPITHELIAL-MESENCHYMAL INDUCTIVE INTERACTIONS.	CC			
CC	-- SUBCELLULAR LOCATION: NUCLEAR (Potential).	CC			
CC	-- DEVELOPMENTAL STAGE: POSTIMPLANTATION.	CC			
CC	-- SIMILARITY: BELONGS TO THE SNAIL FAMILY OF ZINC FINGER PROTEINS.	CC			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	CC			
CC		CC			
DR	EMBL: M95604; AAA03481.1; -	DR			
DR	EMBL: X67253; CAA47075.1; -	DR			
DR	EMBL: U93961; AAB50545.1; -	DR			

DR TRANSFAC; T023333; -;
 DR MGD; NCBI: 98330; Sna;
 DR InterPro; IPR000022; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2_4.
 DR PRINTS; PR00048; ZINCINGER.
 DR PRODOM; PD000003; Znf_C2H2_1.
 DR SMART; SM00355; znf_C2H2_3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 4.
 KW Developmental protein; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat.
 FT DOMAIN 156 259 ZINC FINGERS.
 FT ZN_FING 156 175 C2H2-TYPE.
 FT ZN_FING 180 202 C2H2-TYPE.
 FT ZN_FING 210 230 C2H2-TYPE (ATYPICAL).
 FT ZN_FING 236 259 D->V (IN REF. 2).
 FT CONFFLICT 219 214 AA; 2910 MW; 52E2051224A18DEB CRC64;
 SQ SEQUENCE 264 AA;
 Query Match 13.7%; Score 80; DB 1; Length 264;
 Best Local Similarity 24.5%; Pred. No. 0.76;
 Matches 46; Conservative 13; Mismatches 37; Indels 92; Gaps 11;
 Qy. 4 YEPMPMCAPHLGLFLTG-----LHPALSLP-----LVVTVAGVMSAT----PK- 42
 Db 31 FQDQYDQH----LLAIPPPENLNPAASLPLWLDLSLVPQVRVAVATLPLPESPKA 85
 Qy 43 -----HGLEQCPAPP-PAVTGTGDSG-----AKETVS 70
 Db 86 VELTSLSDEDSGKSSQQPSPPSPAPSFSSTASSELEAFIAFPGLGOLPKQLARLSVA 145
 Qy 71 QKRS-----QSH-----TWCTLAIRPHWITWGVHLRHNS 101
 Db 146 KDPQRKLFNCKYCNKEYLISLGALKHRSHTLPCVCCTGKAESRPMW-LQHVRTHG 204
 Qy 102 ---SASH 105
 Db 205 EKPFCSH 212

RESULT 2

SOX7_HUMAN
 ID SOX7_HUMAN STANDARD; PRT; 388 AA.
 AC Q9BT81;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE transcription factor SOX-7.

GN SOX7_HUMAN
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1] NCBI_TAXID=9606;

RP SEQUENCE FROM N_A.

RC TISSUE=Thymus;
 RX MEDLINE=21350126; PubMed=11691915;
 RA Takash W., Canizares J., Bonneau N., Poulat F., Mattei M.G., Jay P.,
 RA Berta P.;
 RT "SOX7 transcription factor: sequence, chromosomal localisation, expression, transactivation and interference with Wnt signalling.",
 RL Nucleic Acids Res. 29:4274-4283(2001).
 RN [2] SEQUENCE FROM N_A.
 RP SEQUENCE FROM N_A.

TISSUE=Lung;
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kaneko K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
 RA Kawabata B., Suzuki Y., Sugano S., Nagahashi K., Masuho Y., Nagai K.,
 RA Isogai T.;

RESULT 3

SOX7_MOUSE
 ID SOX7_MOUSE STANDARD; PRT; 380 AA.
 AC P40646; Q9R1T6;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE transcription factor SOX-7 (mSOX7).

GN SOX7_MOUSE
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognatheti; Muridae; Murinae; Mus.
 RN [1] NCBI_TAXID=10090;

RP SEQUENCE FROM N_A.

RX MEDLINE=99255428; PubMed=103320775;
 RA Taniguchi K., Hirayama Y., Ogawa M., Sakai Y., Kido S., Aiso S.;
 RT "Isolation and characterization of a mouse SRY-related cDNA, mSox7.",
 RL Biochim. Biophys. Acta 1445:225-231(1999).
 RN [2] SEQUENCE OF 56-109 FROM N_A.

RC STRAIN=Parkes; TISSUE=Brain, and Testis;
 RX MEDLINE=92310993; PubMed=1614875;
 RA Denny P., Swift S., Brand N., Dabholkar N., Barton P., Ashworth A.;
 RT "A conserved family of genes related to the testis determining gene, SRY",
 RL Nucleic Acids Res. 20:2887-2887(1992).
 RC Denny P., Swift S., Brand N., Dabholkar N., Barton P., Ashworth A.;
 CC -I FUNCTION: Transcriptional repressor. Binds to the DNA sequence 5'-
 CC AACAT-3;
 CC -I SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -I TISSUE SPECIFICITY: Ovary and heart.
 CC -I SIMILARITY: Contains 1 HMG BOX.

DE protein 4) (SAPAP4), (PSD-95/SAP90 binding protein 4).
 GN DIGAP OR DAPA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1] SEQUENCE FROM N.A.
 RP
 TISSUE=Brain;
 RX MEDLINE=97277335; PubMed=9115257;
 RA Takeuchi M., Hata Y., Hirao K., Toyoda A., Irie M., Takai Y.;
 RT "SAPPs. A family of PSD-95/SAP90-associated proteins localized at
 postsynaptic density." J. Biol. Chem. 272(11):1943-1951(1997).
 RL [2] FUNCTION: May play a role in the molecular organization of
 synapses and neuronal cell signaling. Could be an adapter protein
 enriching ion channel to the subsynaptic cytoskeleton. May induce
 -I- SUBUNIT: Interacts with DIG1 and DIG4/PSD-95 (By similarity).
 CC SUBCELLULAR LOCATION: Membrane-associated (By similarity).
 CC -I- TISSUE SPECIFICITY: Expressed in brain.
 CC -I- SIMILARITY: BELONGS TO THE SAPAP FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.ebi.sib.ch/announce/>
 or send an email to license@ebi.sib.ch).
 CC
 DR EMBL; U61140; RAB48590.1; -.
 DR InterPro; IPR005026; GKAP.
 DR Pfam; PF03359; GKAP; 1.
 KW Membrane.
 FT DOMAIN 266 273 POLY-PRO.
 SQ SEQUENCE 992 AA; 108034 MW; 8875B67C3FD71F04 CRC64;
 Query Match 12.3%; Score 71.5; DB 1; Length 992;
 Best Local Similarity 31.6%; Pred. No. 20;
 Matches 25; Conservative 6; Mismatches 39; Indels 9; Gaps 3;
 Qy 15 GFLFLGLHPLASLPLIVV----TVAGVMSATPKH-GLEDCPPAPPVAVTGFFGSGAKET 68
 Db 225 GLMTLGROPERNTQPRYFMHAYNTISGHMLKTKNTMTELAPPPIAPPACPSLGVGTD 284
 Qy 69 VSQDKRSQGHTWCTLALPH 87
 Db 285 TNYVKRG---SWSTLFLSH 300
 DB
 RESULT 8
 NTC4_MOUSE
 ID NTC4_MOUSE STANDARD; PRT; 1964 AA.
 AC P31695; Q62389; Q63390; O35442; Q8RIW9; O88314; O88316; Q9R1X0;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neurogenin locus notch homolog protein 4 precursor (Notch 4)
 DE Contains: Transforming protein int-3].
 GN NOTCH4 OR INT3 OR INTP-3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathii; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1] SEQUENCE FROM N.A.
 RX MEDLINE=22194507; PubMed=1312643;
 RA Robbins J., Blondel B.J., Gallahan D., Callahan R.;
 RT "Mouse mammary tumor gene int-3: a member of the notch gene family
 transforms mammary epithelial cells." J. Virol. 66:2594-2599(1992).
 RN [2]
 REVISIONS, SEQUENCE FROM N.A.
 RP MEDLINE=97294599; PubMed=9150355;
 RX
 RA Gallahan D., Callahan R.;
 RI "The mouse mammary tumor associated gene INT3 is a unique member of
 the NOTCH gene family (NOTCH4).";
 RT Oncogene 14:1883-1890(1997).
 RN [3] SEQUENCE FROM N.A.
 RP TISSUE=Lung, and Testis;
 RX MEDLINE=95281668; PubMed=8681805;
 RA Uttyendaele H., Marazzi G., Wu G., Yan Q., Sasoon D., Kitajewski J.;
 RT "Notch4/int-3, a mammary proto-oncogene, is an endothelial
 cell-specific mammalian Notch gene." Development 122:2251-2259(1996).
 RN [4] SEQUENCE FROM N.A.
 RP Rowen L., Mahairas G., Qin S., Ahern M.E., Dankers C., Lasky S., Hood L.;
 RA Lorentz C., Schmidt S., Tipton S., Tralocoff R., Zackrone K., Hood L.;
 RT "Sequence of the mouse major histocompatibility locus class III
 region." Submitted (OCT-1997) to the EMBL/GenBank/DDJB databases.
 RN [5] SEQUENCE OF 1436-1600 FROM N.A.
 RP MEDLINE=99252212; PubMed=10223982;
 RA Lee J.-S., Haruna T., Ishimoto A., Honjo T., Yamagawa S.-I.;
 RT "Intracisternal type A particle-mediated activation of the Notch4/int3
 gene in a mouse mammary tumor: generation of truncated Notch4/int3
 mRNAs by retroviral splicing events." J. Virol. 73:5166-5171(1999).
 RN [6] FUNCTION.
 RP MEDLINE=21244657; PubMed=11344305;
 RA Uttyendaele H., Ho J., Rossant J., Kitajewski J.;
 RT "Vascular patterning defects associated with expression of activated
 Notch4 in embryonic endothelium." Proc. Natl. Acad. Sci. U.S.A. 98:5643-5648(2001).
 RN [7] SEQUENCE OF 1463-1964, POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS
 RP MEDLINE=21523956; PubMed=11518718;
 RA Saxena M.H., Schroeter E.H., Mumm J.S., Kopan R.;
 RT "Murine notch homolog (NL-4) undergoes presenilin-dependent
 proteolysis." J. Biol. Chem. 276:40268-40273(2001).
 RN [8] POST-TRANSLATIONAL PROCESSING.
 RP MEDLINE=21374376; PubMed=11459941;
 RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;
 RT "Conservation of the biochemical mechanisms of signal transduction
 among mammalian Notch family members." Proc. Natl. Acad. Sci. U.S.A. 98:3026-3031(2001).
 CC -I- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Delta1 to regulate cell-fate determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (By similarity). May regulate branching
 CC morphogenesis in the developing vascular system.
 CC -I- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
 CC terminal fragment N(EC) which are probably linked by disulfide
 CC bonds.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus.
 CC -I- TISSUE SPECIFICITY: Highly expressed in lung, moderately in heart
 CC kidney, and at lower levels in the ovary and skeletal muscle. A
 CC very low expression is seen in the brain, intestine, liver and
 CC testis.
 CC -I- DEVELOPMENTAL STAGE: Highly expressed in endothelial cells during
 CC embryonic development from 9.0 d.p.c.
 CC -I- PPM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 trans-Golgi network before it reaches the plasma membrane to yield

CC an active, ligand-accessible form. Cleavage results in a C-terminal fragment N(TM) and a N-terminal fragment N(BC). Following ligand binding, it is cleaved by TNF-alpha converting enzyme (NACE) to yield a membrane-associated intermediate fragment called notch extracellular truncation (NEXT). This fragment is then cleaved by presenilin dependent gamma-secretase to release a notch-derived peptide containing the intracellular domain (NICD) from the membrane.

CC -!- PTM: Phosphorylated.

CC -!- DISEASE: Loss of the extracellular domain causes constitutive activation of the Notch protein, which leads to hyperproliferation of glandular epithelial tissues and development of mammary carcinomas

CC -!- SIMILARITY: BELONGS TO THE NOTCH FAMILY.

CC -!- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.

CC -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.

CC -!- SIMILARITY: CONTAINS 5 ANK REPEATS.

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ish-sib.ch/announce/> or send an email to license@ish-sib.ch).

CC

CC M80456; AAB38377.1; -.

CC EMBL; U43691; AAC52630.1; -.

CC EMBL; AF030001; ABB2004.1; -.

CC EMBL; AB016771; BAA32281.1; ALT SEQ.

DR EMBL; AB016772; BAA32283.1; ALT INIT.

DR EMBL; AB016773; BAA32284.1; ALI INIT.

DR EMBL; AB016774; BAA32285.1; -.

DR PTR; A38072; TWM73.

DR RSPB; P08709; 1B9.

DR MGDB; MGJ:107471; Notch4.

DR InterPro; IPR002110; ANK.

DR InterPro; IPR001512; ASX_hydroxyl.

DR InterPro; IPR000561; EGF-LIKE.

DR InterPro; IPR000742; EGF_2.

DR InterPro; IPR001881; EGF_Ca.

DR InterPro; IPR001438; EGF_II.

DR InterPro; IPR000080; Notch.

DR Pfam; PF00008; EGF; 27.

DR Pfam; PF00033; anky; 6.

DR Pfam; PF00066; notch; 2.

DR PRINTS; PR01415; ANKYRIN.

DR PRINTS; PR00010; EGF_BLOOD.

DR SMART; SM01452; NOTCH.

DR SMART; SM00248; ANK; 5.

DR SMART; SM00179; EGF_Ca; 11.

DR SMART; SM00004; NL; 2.

DR PROSITE; PS00086; ANK_REPEAT; 5.

DR PROSITE; PS0297; ANK REP REGION; 1.

DR PROSITE; PS00010; ASX_HYDROXYL; 11.

DR PROSITE; PS00022; EGF_1; 28.

DR PROSITE; PS01186; EGF_2; 21.

DR PROSITE; PS01187; EGF_Ca; 9.

KW Receptor; Transcription regulation; Activator; Differentiation; developmental protein; Repeat; ANK repeat; EGF-Like domain; transmembrane; Glycoprotein; Signal; Phosphorylation; Proto-oncogene.

FT SIGNAL 1 20 NEUROGENIC.

FT CHAIN 21 1964 TRANSFORMING PROTEIN INN-3.

FT CHAIN 1411 1964 NOTCH EXTRACELLULAR TRUNCATION.

FT CHAIN 1428 1964 NOTCH EXTRACELLULAR DOMAIN.

FT DOMAIN 21 1443 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1444 1964 CYTOSOLIC (POTENTIAL).

FT DOMAIN 21 160 EGF-LIKE 1.

FT DOMAIN 61 112 EGF-LIKE 2.

CC

FT DOMAIN 115 152 EGF-LIKE 3.

FT DOMAIN 153 189 EGF-LIKE 4.

FT DOMAIN 191 229 CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 231 271 EGP-LIKE 5.

FT DOMAIN 273 309 EGP-LIKE 6.

FT DOMAIN 311 350 EGP-LIKE 7, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 352 388 EGP-LIKE 8, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 389 427 EGP-LIKE 9, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 429 470 EGP-LIKE 10.

FT DOMAIN 472 508 EGP-LIKE 11, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 510 546 EGP-LIKE 12, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 548 584 EGP-LIKE 13, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 586 622 EGP-LIKE 14, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 623 656 EGP-LIKE 15, CALCIUM-BINDING (POTENTIAL).

FT DOMAIN 658 686 EGP-LIKE 16.

FT DOMAIN 688 724 EGP-LIKE 17.

FT DOMAIN 726 762 EGP-LIKE 18.

FT DOMAIN 764 800 EGP-LIKE 19.

FT DOMAIN 803 839 EGP-LIKE 20.

FT DOMAIN 841 877 EGP-LIKE 21.

FT DOMAIN 878 924 EGP-LIKE 22.

FT DOMAIN 926 962 EGP-LIKE 23.

FT DOMAIN 964 1000 EGP-LIKE 24.

FT DOMAIN 1002 1040 EGP-LIKE 25.

FT DOMAIN 1042 1081 EGP-LIKE 26.

FT DOMAIN 1083 1122 EGP-LIKE 27.

FT DOMAIN 1126 1167 EGP-LIKE 28.

FT REPEAT 1168 1208 EGP-LIKE 29.

FT REPEAT 1209 1242 LIN_NOTCH 1.

FT REPEAT 1243 1282 LIN_NOTCH 2.

FT REPEAT 1288 1657 LIN_NOTCH 3.

FT REPEAT 1628 1691 ANK 1.

FT REPEAT 1661 1691 ANK 2.

FT REPEAT 1695 1724 ANK 3.

FT REPEAT 1728 1757 ANK 4.

Query Match 12.2% Score 71; DB 1; Length 1964;

Best Local Similarity 24.8%; Pred. No. 45; Mismatches 36; Indels 36; Gaps 5;

Matches 25; Conservative 4; Pairs 36; Gaps 5;

QY 3 GYSEPMPOAHPLGLFLGLHPLASLPLVTVAGMSATPKHGLE---QCPAPP----- 53

Db 1129 GCGPPSPCICH-----NCTCTEPGLGNGFCQTCPDPSPGRCQ 1167

QY 54 -PAVTGFTGDSGAKETVSDKRSQHTW---CTLALPHPW 89

Db 1168 RPGASGCEG-RGGDGTCAGCSGFGDDWDGGDCSLGVBDPW 1207

RESULT 9

EL3B_TRIHA ID EL3B_TRIHA STANDARD; PRIM; 762 AA.

AC P35626; DT 01-OCT-1996 (Rel. 34, Created)

DT 16-Oct-2001 (Rel. 34, Last sequence update)

DE Glucan endo-1,3-beta-D-glucosidase BGN13.1 precursor (EC 3.2.1.39) ((1->3)-beta-D-glucan endohydrolase BGN13.1) ((1->3)-beta-D-glucanase BGN13.1) (Basic Beta-1,3-endoglucanase BGN13.1).

GN BGN13.1.

OS Trichoderma harzianum.

OC Basidiomycota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreales; mitosporic Hypocreales; Trichodermatales; NCBITaxID:5544;

RN [1]

RP SEQUENCE FROM N A., AND PARTIAL SEQUENCE.

RC SPRAINECCT 2413;

RK MEDLINE96074325; Published-7192488;

RA de la Cruz J., Pintor-Toro J.A., Benitez T., Llobell A., Romero L.C.;

RT "A novel endo-beta-1,3-glucanase, BGN13.1, involved in the microparasitism of trichoderma harzianum.";

RL J. Bacteriol. 177:6937-6941 (1995).

CC -!- FUNCTION: INVOLVED IN MYCOPARASITISM, HYDROLYZES YEAST AND FUNGAL CELL WALLS. CLASSIFIED AS A SMALL-OLIGOSACCHARIDE-PRODUCING TYPE

BASED ITS THE END PRODUCTS: GLUCOSE, LAMNARIBIOS OR
 CC LAMNARITRAOSE. OPTIMUM TEMPERATURE FOR ACTIVITY IS 40 DEGREES
 CC CELSIUS, AND INHIBITION OCCURS AT 55 DEGREES CELSIUS.
 CC -I- CATALYTIC ACTIVITY: hydrolysis of 1,3-beta-D-glucosidic linkages
 CC in 1,3-beta-D-glucans.
 CC -I- ENZYME REGULATION: INHIBITED BY GLUCOSE.
 CC -I- SUBCELLULAR LOCATION: Secreted.
 CC -I- DOMAIN: THE C-TERMINAL CYSTEINE-RICH REGION MAY FUNCTION AS A
 CC FUNGAL CELL WALL BINDING DOMAIN.
 CC -I- PTM: DOES NOT SEEM TO BE GLYCOSYLATED.
 CC -I- SIMILARITY: BELONGS TO FAMILY 55 OF GLYCOSYL HYDROLASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; x84085; CAAS5889; 1; "-".
 DR KW Hydrolase; Glycosidase; Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT PROPEP 17 33
 FT CHAIN 34 762 GLUCAN ENDO-1,3-BETA-GLUCOSIDASE BGN13.1.
 FT DOMAIN 520 762 CYS-RICH.
 FT CONFLICT 40 40 P->F (IN REF. 1; AA SEQUENCE).
 FT CONFLICT 759 759 R->P (IN REF. 1; AA SEQUENCE).
 FT SEQUENCE 762 AA: 81246 MW: 0QC25C3C10897AFL CRC64;

Query Match 11.9%; Score 69; DB 1; Length :62;
 Best Local Similarity 32.9%; Pred. No. 26; Length :62;
 Matches 24; Conservative 8; Mismatches 29; Indels 12; Gaps 3;
 AC 094855;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein transport protein Sec24D (SEC24-related protein D).
 GN SEC24D OR KIAA0755.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TAXID=9606;
 RN [1] RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=9907487; PubMed=9872452;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
 RA Tanaka A., Kottani H., Nomura N., Ohara O.;
 RT "prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";
 RT DNA Res. 5:277-286(1998).
 RL [2] RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=9922157; PubMed=10329445;
 RA Tang B.L., Kausalya J., Low D.Y.H., Lock M.L., Hong W.;
 RA "A family of mammalian proteins homologous to yeast Sec24p.";
 RT Biochem. Biophys. Res. Commun. 258:679-684(1999).
 RL DNA Res. 5:277-286(1998).
 RN [1] RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=9907487; PubMed=9872452;
 RA Takebayashi H., Yoshida S., Sugimori M., Kosako H., Kominami R., Nakafuku M., Nakanishi Y.;
 RA "Dynamic expression of basic helix-loop-helix Olig family members: implication of Olig2 in neuron and oligodendrocyte differentiation and identification of a new member, Olig3.";
 RT Mech. Dev. 99:143-148(2000).
 RL [2] RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20182808; PubMed=10719889;
 RA Zhou Q., Wang S., Anderson D.J.;
 RT "Identification of a novel family of oligodendrocyte lineage-specific basic helix-loop-helix transcription factors.";
 RT Neuron 25:331-343(2000).
 CC -I- FUNCTION: COMPONENT OF THE COPII COAT, THAT COVERS ER-DERIVED
 CC VESICLES INVOLVED IN TRANSPORT FROM THE ENDOPLASMIC RETICULUM TO
 CC THE GOLGI APPARATUS. COPI ACTS IN THE CYTOPLASM TO PROMOTE
 CC TRANSPORT OF SECRETORY, PLASMA MEMBRANE, AND VACUOLAR PROTEINS
 CC FROM THE ENDOPLASMIC RETICULUM TO THE GOLGI COMPLEX.
 CC -I- SUBUNIT: COPI IS COMPOSED OF AT LEAST FIVE PROTEINS: THE SEC23/24
 CC COMPLEX, THE SEC13/31 COMPLEX AND SAIL.
 CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC AND PERNUCLEAR (BY SIMILARITY).
 CC PLACENTA, PANCREAS, HEART AND LIVER.
 CC -I- SIMILARITY: BELONGS TO THE SEC23/SEC24 FAMILY. SEC24 SUBFAMILY.
 DR EMBL; AB018298; BAA34475; 1; "-".
 DR EMBL; AFL3464; AAD28756; 2; "-".
 DR Genew; HGNC:10706; SGC24D.
 KW Transport; Protein transport; Golgi stack; Endoplasmic reticulum;
 KW Multigene family.
 FT DOMAIN 363 398 ZINC FINGER-LIKE.
 FT DOMAIN 10 293 PRO-RICH.
 FT SEQUENCE 1032 AA: 112999 MW: 1EA80215EB979A7 CRC64;

Query Match 11.9%; Score 69; DB 1; Length 1032;
 Best Local Similarity 28.4%; Pred. No. 36; Length 1032;
 Matches 21; Conservative 7; Mismatches 30; Indels 16; Gaps 2;
 AC 094855;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Oligodendrocyte lineage transcription factor 2.
 GN OLIG2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TAXID=10090;
 RN [1] RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20344962; PubMed=11091082;
 RA Nakafuku M., Nakanishi Y.;
 RA "Dynamic expression of basic helix-loop-helix Olig family members: implication of Olig2 in neuron and oligodendrocyte differentiation and identification of a new member, Olig3.";
 RT Mech. Dev. 99:143-148(2000).
 RL [2] RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20182808; PubMed=10719889;
 RA Zhou Q., Wang S., Anderson D.J.;
 RT "Identification of a novel family of oligodendrocyte lineage-specific basic helix-loop-helix transcription factors.";
 RT Neuron 25:331-343(2000).
 CC -I- FUNCTION: Required for oligodendrocyte and motor neuron

specification in the spinal cord.

-1- SUBCELLULAR LOCATION: Nuclear (Potential).

-1- TISSUE SPECIFICITY: Specifically expressed in zones of neuroepithelium from which oligodendrocyte precursors emerge, as well as in the precursors themselves.

-1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHHL) FAMILY OF TRANSCRIPTION FACTORS. "ATONAL" SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AR232929; AAF61722.1; -

MGB; MG1:153531; OL192.

DR InterPro; IPR01052; HHL_basic.

PFAM; PF00100; HHL_1.

SMART; SM0053; HHL_1.

DR PROSITE; PS00038; HHL_1; 1.

DR PROSITE; PS50888; HHL_2; 1.

KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN.

FT DNA-BIND 109 120

FT DOMAIN 121 162 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).

FT DOMAIN 77 88 POLY-SER.

FT DOMAIN 206 212 POLY-ALA.

FT DOMAIN 225 237 POLY-ALA.

FT DOMAIN 266 270 POLY-GLY.

FT DOMAIN 273 284 POLY-GLY.

FT CONFLICT 198 L -> I (IN REF. 2).

FT CONFLICT 205 P -> PTHRQAP (IN REF. 2).

FT SEQUENCE 323 AA: 32406 MW: 598AE76CB512D716 CRC64;

Query Match 11.8%; Score 69.5; DB 1; Length 323;

Best Local Similarity 23.9%; Pred. No. 12; Mismatches 33; Conservative 13; Matches 35; Indels 57; Gaps 7;

Qy 2 LGYSEMP-CAHPLGLFLGLHPLSLPLV-----VTVAGVMSAT----- 40

Db 192 LAMSAPLTAPTAAPAAAHAAHPAVHPLPPAAAAAAAAAASASLPGSGLSSV 251

Qy 41 ---PKHGLEQCPAPPAPPVAT---GFMGDSAKETVSQDKRSQGH-----TWCTLA 84

Db 252 GSIRRPHGLKFSAAAAPLGGGGGGSGGGFQ-----HWGGMPCPSCMCOV 301

Qy 85 LPFWLWTWGLRNHVSS 102

Db 302 PPH-----HHVSA 309

RESULT 12

MTD1_ARATH STANDARD; PRT; 357 AA.

ID Q02971; Q9S4J9; [1]

AC 001-NOV-1995 (Rel. 32, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-oct-2001 (Rel. 40, Last annotation update)

DE Probable mannitol dehydrogenase 1 (EC 1.1.1.255) (NAD-dependent DE mannitol dehydrogenase 1).

GN Arabidopsis thaliana (Mouse-ear cress).

OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; eudicots; Rosidae; eu rosids II; Brassicales; Brassicaceae; Arabidopsidae; Arabidopsis; NCBI_TAXID=3702;

RN ELI3-1 OR AR232929 OR F20010.100.

SPQUENCE FROM N.A.

RP SPROT-NCBI_COLUMBIA;

RX MEDLINE=93009840; PubMed=1464303;

RA Kiedrowski S., Kawalleck P., Hailbrock K., Sonnichsen J.E., Dangl J.L.;

RA DR EMBL; AL03538; CAB37538.1; -

DR EMBL; AL161592; CAB80463.1; -

RT "Rapid activation of a novel plant defense gene is strictly dependent on the Arabidopsis RPM1 disease resistance locus.";

RT EMBO J. 11:4677-4684(1992).

[2]

SEQUENCE FROM N.A.

RC STRAIN=CV. Columbia;

RC MEDLINE=20083488; PubMed=16617198;

RA Mayer K.F.X., Schueler C., Wambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N., Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M., Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M., Kreis M., Deisenby P., Puigdomenech P., Watson M., Schmidtheinl T., Reichert B., Portereille D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hoheisel J., Zimmermann W., Wedler H., Ridley P., Langham S.A., McCullagh B., Bilham L., Robben J., Vandenhussche F., Van der Schueren J., Grymonprez B., Chuang Y.-J., De Clercq R., Braeken M., Weltens I., Vett M., Bastiaens I., Aert R., Defoor E., Weltzengger T., Bothe G., Ramsperger U., Hilpert H., Braun M., Holzer E., Brandt A., Peters S., van Staerwen M., Dirkse W., Moijman P., Klein Laakhorst R., Rose M., Hauf J., Koetter P., Bernreiter S., Hempel S., Feldpausch M., Lambeth S., van den Daele H., De Keyser A., Buxshaert C., Gielen J.J., Villaruel R., De Clercq R., Van Montagu M., Rogers J., Cronin A., Quall M., Bray-Alien S., Clark L., Doogert J., Hall S., Kay M., Leonard N., McLay K., Mayes R., Pettet A., Rajanadean M.A., Lyne M., Benes V., Rechmann S., RA Borikova D., Bloecker H., Scharfe M., Grimm M., Loehnert T.-H., Dose S., de Haan M., Maarse A.-C., Schaefer M., Mueller-Neuer S., Gabel C., Fuchs M., Fartmann B., Grandérath K., Dauner D., Herzl A., Neumann S., Argiriou A., Vitale D., Liuguo R., Piravandi E., Massenet O., Quiley F., Clabaugh G., Muendlein A., Feilzer R., Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S., RA Chef d'Or F., Cooke R., Berger C., Monfort A., Casacuberta E., RA Gibbons T., Weber N., Vandembol M., Bargues M., Terol J., Torres A., Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T., Heijnen L., Schwarz S., Schollier P., Heber S., Francs P., Blelkie C., RA Frishman D., Haase D., Lemcke K., Meyers H.-W., Stocker S., Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K., RA Parneil L., Dedha N., Gnoj L., Schutz K., Huang E., Spiegel L., RA Sethon M., Murray J., Sheet P., Cordes M., Abu-Threideh J., RA Stokking T., Kajicki J., Graves T., Harmon G., Edwards J., RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D., RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kramer J., Fulton L., Mardis E., Date M., Peipin K., Hillier L., RA Nelson J., Spieth J., Ryan E., Andrews S., Geissel C., Layman D., RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshu C., RA Antoniou B., Zidani M., Strong C., Lannai B., Yordan C., RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R., RA Swayb T.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S., RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lohdi M., Johnson A., RA Chen E., Marra M., Martienssen R., McCombie W.R.; RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana."; Nature 402:769-777(1999).

CC -1- FUNCTION: OXIDIZES MANNITOL TO MANNOSE. PROVIDES THE INITIAL STEP BY WHICH TRANSLOCATED MANNITOL IS COMMITTED TO CENTRAL METABOLISM AND, BY REGULATING MANNITOL POOL SIZE, IS IMPORTANT IN REGULATING SALT TOLERANCE AT THE CELLULAR LEVEL (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: D-mannitol + NAD(+) = D-mannose + NADH.

CC -1- COFACTOR: ZINC (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR X67816; CAA48027.1; -

DR EMBL; AL03538; CAB37538.1; -

DR EMBL; AL161592; CAB80463.1; -

DR	PIR; S28044; S28044;	RT	"Molecular cloning of chicken aggrecan. Structural analyses.";
DR	InterPro; IPR00328; ADH_zinc;	RL	Biochem. J. 288:903-910(1992).
DR	InterPro; IPR002085; Adh_zn_family.	RN	[4]
DR	Pfam; PF00107; adh_zinc_1;	RP	ERRATUM.
KW	Oxidoreductase; NAD; ADH_zinc; 1.	RX	MDBLINK=94107258; PubMed=1280087;
FT	METAL	RA	Chandrasekaran L., Tanzer M.L.;
FT	METAL	RL	Biochem. J. 296:885-887(1993).
FT	METAL	RN	[5]
FT	METAL	RP	SEQUENCE OF 1493-1610 FROM N.A.
FT	METAL	RC	STRAIN=White Leghorn; TISSUE=Chondrocytes;
FT	METAL	RX	MDBLINK=95128519; PubMed=827752;
FT	METAL	RA	Primorac D., Stover M.L., Clark S.H., Rose D.W.;
FT	METAL	RT	"Molecular basis of nanomelia, a heritable chondrodystrophy of
FT	METAL	RT	chicken.";
FT	CONFFLICT	RL	Matrix Biol. 14:297-305(1994).
FT	CONFFLICT	RN	[6]
FT	SEQUENCE	RP	SEQUENCE OF 1894-2109 FROM N.A.
SO	SEQUENCE	RX	MEDLINE=89008500; PubMed=3170613;
Query	Match	RA	Tanaka T., Har-El R., Tanzer M.L.;
Best	Local Similarity	RT	"Partial structure of the gene for chicken cartilage proteoglycan
Matches	25.2%;	RT	core protein.>";
35;	Conservative	J. Biol. Chem. 263:15831-15835(1988).	
Qy	3	RN	[7]
Db	333 GYSEPMPCAHPLGLFLGLPAISL---PLVVTAGVMSATPKHGLEQCPAPPATG	RP	SEQUENCE OF 1693-1855 AND 1893-2109 FROM N.A.
	58	RX	MEDLINE=86259736; PubMed=3460082;
Qy	59	RA	Sai S., Tanaka T., Kosher R.A., Tanzer M.L.;
Db	-----FGFDGAKETV---SODKRSQHTWC-----	RT	"Cloning and sequence analysis of a partial cDNA for chicken
	81	RT	cartilage proteoglycan core protein.>";
Qy	188 LGGLGHVAVKAMGRVTVTISERRDDEAVTRLGADAFALVSRDPQKMDAMGTMGI	RL	Proc. Natl. Acad. Sci. U.S.A. 83:5051-5055(1986).
Db	248 IDTVSATPHPLPULGILKN	CC	-I- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
	266	CC	MATRIX OF CARTILAGENOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
		CC	IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVOIDLY TO
		CC	Hyaluronic acid via an amino-terminal globular region. MAY PLAY A
		CC	REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
		CC	-I- SUBCELLULAR LOCATION: Secreted; extracellular matrix (BY
		CC	similarity);
		CC	-I- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
		CC	TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
		CC	MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
		CC	CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
		CC	THE A, B, C MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
		CC	AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
		CC	AND G3.
		CC	-I- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
		CC	CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.
		CC	-I- DISEASE: DEFECTS IN AGC1 ARE THE CAUSE OF NANOMELIA, A LETHAL
		CC	CONNECTIVE TISSUE DISORDER AFFECTING CARTILAGE DEVELOPMENT
		CC	(CHONDRODYSPIPHY) CHARACTERIZED BY SHORTENED AND MALFORMED LIMBS.
		CC	AGC1 IS TRUNCATED AT ITS C-TERMINAL IN THE CS-2 BINDING DOMAIN
		CC	AND IS NOT ANYMORE SECRETED FROM THE CHONDROCYTES.
		CC	-I- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
		CC	-I- SIMILARITY: CONTAINS 4 LINK DOMAINS.
		CC	-I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
		CC	-I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
		CC	-I- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
		CC	-I- SIMILARITY: BELONGS TO THE AGGREGAN/VERSGAN PROTEOGLYCAN FAMILY.
		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation
		CC	the European Bioinformatics Institute. There are no restrictions on its
		CC	use by non-profit institutions as long as its content is in no way
		CC	modified and this statement is not removed. Usage by and for commercial
		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
		CC	-----
		CC	DR EMBL; M38187; AA448731; -;
		CC	DR EMBL; M88101; ; NOT_ANNOTATED_CDS.
		CC	DR EMBL; S74657; AAC60751; -;
		CC	DR EMBL; S74656; AAC60751; JOINED.
		CC	DR EMBL; J04028; AA448719; -;
		CC	DR EMBL; M13993; AA448720; -;
		CC	DR EMBL; I21913; AA319128; -;
		CC	DR EMBL; M38187; AA448731; -;
		CC	DR EMBL; S74657; AAC60751; -;
		CC	DR EMBL; S74656; AAC60751; JOINED.
		CC	DR EMBL; J04028; AA448719; -;
		CC	DR EMBL; M13993; AA448720; -;
		CC	DR EMBL; A25442; A25442.

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu B., Dreano S., Gloux S., Delaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt K., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Taillade V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potschkin J.,
 RA Spakowski G.V., Usery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: PLAYS A ROLE IN THE CELL DIVISION RING; INVOLVED IN CYTOKINESIS.
 CC -!- COMPOSITION: OF THE CELL DIVISION RING; INTERACTS WITH PROFILIN.
 CC -!- SIMILARITY: TO YEAST BNL1.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -!- FUNCTION: PLAYS A ROLE IN THE CELL DIVISION RING; INTERACTS WITH PROFILIN.

CC -!- SIMILARITY: TO YEAST BNL1.

DR EMBL: AF317653; AAK12638 1;

DR Gene; HGNC:13301; GPR62.

DR MIM: 606917; -.

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7m_1; 1.

DR PRINTS; PRO0237; GPCR_RHODOPSN.

DR PROSITE; PS00237; G-PROTEIN_RECER_F1_1; FALSE_NEG.

DR PROSITE; PS50362; G-PROTEIN_RECER_FL_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein.

FT DOMAIN 1 18 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 19 39 1 (POTENTIAL).

FT DOMAIN 40 53 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 54 74 2 (POTENTIAL).

FT DOMAIN 75 91 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 92 112 3 (POTENTIAL).

FT DOMAIN 113 129 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 130 150 4 (POTENTIAL).

FT DOMAIN 151 177 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 178 198 5 (POTENTIAL).

FT DOMAIN 199 239 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 240 260 6 (POTENTIAL).

FT DOMAIN 261 272 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 273 293 7 (POTENTIAL).

FT DOMAIN 294 368 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 299 325 ALA-RICH.

FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT SEQUENCE 368 AA; 37628 MW; 9CFF9529BD12C75 CRC64;

Query Match 11.4%; Score 66.5; DB 1; Length 1841;
 Best local Similarity 33.3%; Pred. No. 1.1e-02;
 Matches 23; Conservative 6; Mismatches 27; Indels 13; Gaps 4;

Qy 7 PMPCAHPGLFLGLPHALPSLVLV--VTVAGVMSATPKHGLEQCPAPPAPPAVNGFTGDS 63
 Db 912 FLPVKTSLNTEF--SHPD-SVNTVANDTSVAGVMRAFPF---PPPPPPPLVSAAGK 961

Qy 64 GAKERVSQD 72
 Db 962 FVSPAVSN 970

RESULT 15

GP62_HUMAN STANDARD; PRT; 368 AA.

ID GP62_HUMAN STANDARD; PRT; 368 AA.

AC 0B2J77; DOMAIN 1 18 EXTRACELLULAR (POTENTIAL).

DT 15-JUN-2002 (Rel. 41, created)
 DT 15-JUN-2002 (Rel. 41, last sequence update)
 DE 15-JUN-2002 (Rel. 41, last annotation update)

DE Probable G protein-coupled receptor GPR62.

GN GPR62.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TAXID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=1105013; PubMed=11165367;

RA Lee D.K., George S.R., Cheng R., Nguyen T., Liu Y., Brown M., Lynch K.R., O'Dowd B.F.;
 RT "Identification of four novel human G protein-coupled receptors expressed in the brain.";
 RL Brain Res. Mol. Brain Res. 86:13-22(2001).

CC -!- FUNCTION: Orphan receptor.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- TISSUE SPECIFICITY: Expressed in brain; detected in the basal forebrain, frontal cortex, caudate, putamen, thalamus and hippocampus.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -!- FUNCTION: PLAYS A ROLE IN THE CELL DIVISION RING; INTERACTS WITH PROFILIN.

CC -!- SIMILARITY: TO YEAST BNL1.

DR EMBL: AF317653; AAK12638 1;

DR Gene; HGNC:13301; GPR62.

DR MIM: 606917; -.

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7m_1; 1.

DR PRINTS; PRO0237; GPCR_RHODOPSN.

DR PROSITE; PS00237; G-PROTEIN_RECER_F1_1; FALSE_NEG.

DR PROSITE; PS50362; G-PROTEIN_RECER_FL_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein.

FT DOMAIN 1 18 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 19 39 1 (POTENTIAL).

FT DOMAIN 40 53 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 54 74 2 (POTENTIAL).

FT DOMAIN 75 91 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 92 112 3 (POTENTIAL).

FT DOMAIN 113 129 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 130 150 4 (POTENTIAL).

FT DOMAIN 151 177 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 178 198 5 (POTENTIAL).

FT DOMAIN 199 239 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 240 260 6 (POTENTIAL).

FT DOMAIN 261 272 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 273 293 7 (POTENTIAL).

FT DOMAIN 294 368 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 299 325 ALA-RICH.

FT CARBOHYD 3 3 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT SEQUENCE 368 AA; 37628 MW; 9CFF9529BD12C75 CRC64;

Query Match 11.3%; Score 66; DB 1; Length 368;
 Best local Similarity 26.9%; Pred. No. 24; Matches 32; Conservative 6; Mismatches 37; Indels 44; Gaps 5;

Qy 7 PMPC-----AHPGLFLGL-----HPLASLPLVTVAGVMS 38
 Db 85 PAPCRARFLSAAIPLACTLQVVAAGLARVLIVHLRPGSSRPPVVLVLTAVWAAGLG 144

Qy 39 ATPKIGLEQCPAPPAA--VIGFGDGSAGETVSQDKRSQGHTWCTLAIPHWMUTWG 94
 Db 145 AHSLG--PPPPAPPAPARCSVLAGGLGPPRL-----WALLAFLPALLLG 190

Search completed: June 17, 2003, 10:35:27
 Job time : 24 secs

THIS PAGE BLANK (USPTO)

Qy 102 ---SASH 105
Db 205 EKPFSCSH 212

RESULT 2

G02409 protein kinase C-binding protein RACK17 - human (fragment)
C;Species: Homo sapiens (man)
C;Accession: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 05-Nov-1999
R;Kuroda, S.; Tokunaga, C.; Kirohara, Y.; Konishi, H.; Kikkawa, U.
R;Reference number: H01212
A;Accession: G02409
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Residues: 1-334 <R0X>
A;Cross-references: EMBL:U48250; NID:91199656; PIDN:ACT72247.1; PID:gi1199657

RESULT 3

Query Match 12.6%; Score 73.5; DB 2; Length 334;
Best Local Similarity 25.5%; Pred. No. 7; 7;
Matches 35; Conservative 12; Mismatches 35; Indels 55; Gaps 7;
Qy 2 LGVSEPMPC--AHPGLFLGLQPLSLPV-----VTVAGVNSAT----- 40
Db 203 LAHSAPLPLPATAHAAAHAAHHPAVHHPILPPAAAAAAAASLPSGLPSV 262
Qy 41 ---PKIGLEQCPAPPAP--VYFTGSGAKENVSQKRSGH-----TWCLAL 85
Db 263 GSTRRPHGKPSAANAAAPLGGGGSGA-----SGGFOHWGSGMPCPCSMCOVPP 313
Qy 86 PHWLTWVGHLRNHSV 102
Db 314 PH-----HHVSA 320

RESULT 4

R86244 hypothetical protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Accession: F86244 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Fedderspiel, N.A.; Kaul, S.; White, O.; Alonso, C.;Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;Ansien, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziali, A.;Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F86244
A;Status: preliminary
A;Residues: prelminary DNA
A;Cross-references: EMBL:AU035636; PIDN:CAB38479.1; GSPDB:GN00070; SCEDB:SCH5.05c
A;Genetics:
A;Gene: SCEDB:SCH5.05c

Query Match 12.4%; Score 72; DB 2; Length 1089;
Best Local Similarity 32.8%; Pred. No. 37; 37;
Matches 39; Conservative 9; Mismatches 45; Indels 26; Gaps 8;
C;Genetics:
C;Map position: 1

Query Match 12.5%; Score 72.5; DB 2; Length 554;
Best Local Similarity 32.4%; Pred. No. 16; 16;
Matches 22; Conservative 3; Mismatches 30; Indels 13; Gaps 1;
Qy 21 IHPALSLPLVVTAGVMSATPKHGLEQCP-----APPAVATGTDGSAKE 67
Db 232 IFRPPPLPMAVRKGVAAPPLPPGTAALPPPLPMAKGVAAPPPPGARGGLGAK 291

Qy 68 TVSDOKIS 75
Db 513 ALEGGDGAGKSTOAELAEWIRKGKHE--VVLTRPAGATPVCKRLRSILLDVSAGSLSH 569
JN0707 proteinaceous alpha-amylase inhibitor precursor - Streptomyces nitrosporeus

Db 292 VTSKLKRS 299
RESULT 5

T11753 mullerian inhibiting factor precursor - pig
C;Species: Sus scrofa domesticus (domestic Pig)
C;Accession: T11753 #sequence_revision 16-Jul-1999 #text_change 26-Aug-1999
R;Daneau, I.; Silversides, D.W.
submitted to the EMBL Data Library, July 1998
A;Reference number: Z17334
A;Accession: T11753
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-575 <DAN>
A;Cross-references: EMBL:U80853; NID:93299818; PID:gi3299819
A;Genetics:
A;Gene: MIS
C;Function:
A;Description: causes regression of the Mullerian duct, inhibits the growth of tumors
ibition
A;Note: Produced by the Sertoli cells of the testis
C;Superfamily: Inhibin
C;Keywords: cytotoxin; glycoprotein; gonadal differentiation; testis
Query Match 12.4%; Score 72; DB 2; Length 575;
Best Local Similarity 29.1%; Pred. No. 19; 19;
Matches 32; Conservative 9; Mismatches 33; Indels 36; Gaps 5;
Qy 1 MGVSSEPMPC-AHPGLFLGLQPLSLPVVWVAGMSATPKHGLEQCPAPPVTFG 59
Db 159 LURFQEPLPGEASPLELLALVLYPGPGEVWTGAGLGAQ----SLCP----- 203
Qy 60 TGSGAKENVSQDKRSGHTWCLALPFLWVHG---LRNHVSSAS 104
Db 204 TRDSG-----FLALAVDPRERAWRGSGLALTIRRNGAS 238

RESULT 6

T36663 protein kinase, transmembrane - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Accession: T36663 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.; submitted to the EMBL Data Library, March 1999
A;Reference number: Z21611
A;Accession: T36663
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: DNA
A;Residues: 1-1089 <SEE>
A;Cross-references: EMBL:AU035636; PIDN:CAB38479.1; GSPDB:GN00070; SCEDB:SCH5.05c
A;Experimental source: strain A3(2)
A;Genetics:
A;Gene: SCEDB:SCH5.05c

C;Species: Streptomyces nitrosporeus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 28-May-1999
C;Accession: JN0707
C;Authors: Cole, S.T.; Brosh, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Sunitani, J.; Kavaguchi, T.; Hattori, N.; Murao, S.; Arai, M.
C;Keywords: Biotechnol. Biochem. 57, 1243-1248, 1993
A;Title: Molecular cloning and expression of proteinaceous alpha-amylase inhibitor gene
A;Reference number: JN0707; MUID:93379351; PMID:7764011
A;Molecule type: DNA
A;Residues: 1-110 <SUM>
A;Cross-references: GB:S65457; NID:9451243; PIDN:AAC60452.1; PID:9451244
C;Comment: This enzyme is active specifically against alpha-amylases from animals but in C;Superfamily: Streptomyces alpha-amylase inhibitor
C;Keywords: alpha-amylase inhibitor
F;1-33/domain: signal sequence #status predicted <SI>
F;34-110/product: proteinaceous alpha-amylase inhibitor #status predicted <MAT>
Query Match 12.2%; Score 71; DB 2; Length 110;
Best Local Similarity 29.9%; Pred. No. 4.3%; Mismatches 44; Indels 10; Gaps 4;
Matches 26; Conservative 7; Mismatches 44; Indels 10; Gaps 4;
C;Accession: 21 LHPAISLPLVTVAGVMSATPKHGLEQCPAPPAPPA-VTGETW-----PSGAKETVSQDK 73
Db 8 ISPANITIATV-AAGLMTAAAPSAAATGAPAPACVSEFQSWRYTDVHNGCSEIVSVTW 66
C;Genetics:
C;Superfamily: Mycobacterium tuberculosis probable polyketide synthase pks1; acyl car
C;Keywords: carrier protein
C;Description: Sequence of the mouse major histocompatibility locus class III region.
A;Reference number: Z16543
A;Accession: T09059
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: DNA
A;Residues: 1-1964 <ROW>
A;Cross-references: EMBL:AF030001; NID:92564945; PID:9:564947
C;Genetics:
A;Gene: notch4
A;Map position: 17
A;Introns: 2/21; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 67
C;Superfamily: unassigned EGFR-related proteins; EGF homology
C;Keywords: receptor; signal transduction
F;154-545/Domain: EGF homology <EGF>
Query Match 12.2%; Score 71; DB 2; Length 1964;
Best Local Similarity 24.8%; Pred. No. 86; Mismatches 25; Conservative 4; Mismatches 36; Indels 36; Gaps 5;
Matches 25; Conservative 4; Mismatches 36; Indels 36; Gaps 5;
C;Accession: 3 GYSEPMPCAHPIGLFLGLHPLSPLVTVAGVMSATPKHGLE---QCPAPP----- 53
Db 1129 GCGPSPSPCLH-----NGCTETPGLGNPQFQCTCPPDSPGPRCQ 1167
C;Accession: A75450
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <PHI>
A;Cross-references: GB:AE001952; GB:AE000513; NID:96458725; PIDN:AAF10584.1; PID:9645
A;Experimental source: strain R1
C;Genetics:
A;Gene: dr109
A;Map Position: 1
Query Match 11.9%; Score 69.5; DB 2; Length 434;
Best Local Similarity 29.3%; Pred. No. 25; Mismatches 27; Conservative 9; Mismatches 37; Indels 19; Gaps 5;
Matches 27; Conservative 9; Mismatches 37; Indels 19; Gaps 5;
C;Accession: G70668
C;Cross-references: GB:AE001952; GB:AE000513; NID:96458725; PIDN:AAF10584.1; PID:9645
C;Keywords: Mycobacterium tuberculosis
C;Species: Mycobacterium tuberculosis
N;Alternate names: polyketide synthase pks002c
C;Species: Mycobacterium tuberculosis

THIS PAGE BLANK (USPTO)

Copyright (c) 1993 - 2003 Compugen Ltd.

GenCore version 5.1.6

ON protein - protein search, using sw model
Run on: June 17, 2003, 10:33:46 ; Search time 25 Seconds
(without alignments)

18.823 Million cell updates/sec

Title: US-09-955-807-2
Perfect score: 582

Sequence: 1 MGYSEPMCAHPLGLFLIG..... PHWMLTWVGHLRNHVSASH 105

Scoring table: BLOSUM62
Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A.COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/6A.COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/6B.COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	71	12.2	1964	4 US-09-467-997-1	Sequence 1, Appl
2	67.5	11.6	1185	4 US-08-664-962B-2	Sequence 2, Appl
3	67.5	11.6	1185	4 US-09-111-743-2	Sequence 3, Appl
4	66	11.3	306	1 US-08-868-852-13	Sequence 4, Appl
5	66	11.3	306	2 US-08-225-940-13	Sequence 5, Appl
6	66	11.3	306	2 US-08-976-838-13	Sequence 6, Appl
7	66	11.3	799	2 US-08-525-940-23	Sequence 7, Appl
8	66	11.3	799	2 US-08-976-830-23	Sequence 8, Appl
9	66	11.3	881	2 US-08-525-940-21	Sequence 9, Appl
10	66	11.3	881	2 US-08-976-838-21	Sequence 10, Appl
11	66	11.3	915	2 US-08-225-940-18	Sequence 11, Appl
12	66	11.3	915	2 US-08-976-838-18	Sequence 12, Appl
13	66	11.3	915	4 US-09-114-55B-2	Sequence 13, Appl
14	66	11.3	915	4 US-09-214-55B-7	Sequence 14, Appl
15	64.5	11.1	326	1 US-08-053-867A-2	Sequence 15, Appl
16	64.5	11.1	326	1 US-08-868-332C-32	Sequence 16, Appl
17	64.5	11.1	2972	4 US-08-979-181-2	Sequence 17, Appl
18	64.5	11.1	3118	4 US-09-079-181-4	Sequence 18, Appl
19	64	11.0	380	2 US-09-026-587-4	Sequence 19, Appl
20	64	11.0	380	2 US-09-227-420-4	Sequence 20, Appl
21	63.5	10.9	369	1 US-07-688-332C-32	Sequence 21, Appl
22	63.5	10.9	369	2 US-08-974-379C-32	Sequence 22, Appl
23	63.5	10.9	369	3 US-09-146-219A-32	Sequence 23, Appl
24	63.5	10.9	369	3 US-08-206-188B-32	Sequence 24, Appl
25	63.5	10.9	369	4 PCT-US91-02714-31	Sequence 25, Appl
26	63.5	10.9	380	4 US-09-673-397-2	Sequence 26, Appl
27	63.5	10.9	380	4 US-09-673-397-6	Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-09-467-997-1
; Sequence 1, Application US/09467997
; Patent No. 6379925
; GENERAL INFORMATION:
; APPLICANT: Kitaewki, Jaan
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
; FILE REFERENCE: 53663-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09-467,997
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ_ID NO 1
; LENGTH: 1964
; TYPE: PRT
; ORGANISM: mouse
US-09-467-997-1

Query Match Best Local Similarity 12.2%; Score 71; DB 4; Length 1964;
Matches 25; Conservative 24.8%; Pred. No. 25; Gaps 36; Indels 36; Gaps 5;

QY 3 GISEPPMCAHPLGLFLIGLHPALSLPLVVTAGMSATPKHGLE---QCPAPP----- 53
Db 1129 GCGPPSPCLH-----NCTCTESPGIGGNPFGQCTCPDPSGPRCQ 1167

QY 54 -PAVTFGFDGSAKETYSQDKRSQGHTW---CTLA1PHW 89
Db 1168 RPGASGCCG-RGGDTCDAGCSGGDPDWGDPGCGSLGVDPW 1207

RESULT 2
US-08-664-962B-2
; Sequence 2, Application US/08664962B
; Patent No. 6218162
; GENERAL INFORMATION:
; APPLICANT: Krystal, Gerald
; TITLE OF INVENTION: SH2-CONTAINING INOSITOL-PHOSPHATASE
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MERCHANT & GOULD
; STREET: 3100 No. 621816-West Center, 90 South Seventh Street
; CITY: Minneapolis
; STATE: Minnesota
; COUNTRY: U.S.A.
; ZIP: 55402-4131
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/664,962B
 FILING DATE: 14-JUN-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Mueller, Douglas P.
 REGISTRATION NUMBER: 30,300
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 612-332-5300
 TELEX/FAX: 612-332-9081
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1185 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-664-962B-2

RESULT 3
 US-08-311-743-2
 ; Sequence 2, Application US/09311743
 ; Patent No. 6238903
 GENERAL INFORMATION:
 APPLICANT: Krystal, Gerald
 TITLE OF INVENTION: SH2-CONTAINING INOSITOL-PHOSPHATASE
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BERSEKIN & PARR
 STREET: 40 King Street West
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5H 3Y2
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/368,852
 FILING DATE: 05-JAN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Verser, Carroll Talkington
 REGISTRATION NUMBER: 37,459
 REFERENCE/DOCKET NUMBER: 2848-11
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 303-863-9700
 TELEX/FAX: 303/863 0223
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 306 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-368-852-13

Query Match 11.6%; Score 67.5; DB 4; Length 1185;
 Best Local Similarity 28.9%; Pred. No. 34; Mismatches 29; Indels 15; Gaps 4;
 Matches 22; conservative 10; Mismatches 29; Indels 15; Gaps 4;

QY 6 EPMPCAHPIGLFLIGL-HPALSLPLVWVAGVMSATPKHGLECP--APPRAVTGFGC 61
 Db 1041 EPPCPDPD----GSSSPSIVLPKAQEVEVKGTS----KQAPPVVLGPTRPRISFTC 1089
 QY 62 DSGAKETVSQDKRSQG 77
 Db 1090 SSSAEGRMTSGDKSQG 1105

RESULT 4
 US-08-368-852-13
 ; Sequence 13, Application US/08368852
 ; Patent No. 5691183
 GENERAL INFORMATION:
 APPLICANT: Franzisoff, Alex
 TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross & McIntosh
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: CO
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/368,852
 FILING DATE: 05-JAN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Verser, Carroll Talkington
 REGISTRATION NUMBER: 37,459
 REFERENCE/DOCKET NUMBER: 2848-11
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 303-863-9700
 TELEX/FAX: 303/863 0223
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 306 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-368-852-13

Query Match 11.3%; Score 66; DB 1; Length 306;
 Best Local Similarity 27.5%; Pred. No. 9.8; Mismatches 26; Indels 24; Gaps 5;
 Matches 25; conservative 16; Mismatches 26; Indels 24; Gaps 5;

QY 37 MSATPKRG---LEQCPAPPRAVTGFGDSSGAKETVSQDKR---SOGHTWCILALP-- 86
 Db 200 ISSITAESGRKPKWYLEEC--SSTIATYSSGEYSYDKRILITDLRQRCTDNHTGTSASAPMA 257
 QY 87 -----HPWLTW--VGHLRNHVSSASH 105
 Db 258 AGITALALEANPELTWDVQHVIVRTSRAGH 288

RESULT 5
 US-08-325-940-13
 ; Sequence 13, Application US/08525940
 ; Patent No. 5866351
 GENERAL INFORMATION:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/976, 838
 FILING DATE: 07-JUL-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2948-11-C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 FAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 13:
 LENGTH: 306 amino acids
 SEQUENCE CHARACTERISTICS:
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-976-838-13

RESULT 7
 US-08-525-940-23
 Query Match 11.3%; Score 66; DB 2; Length 306;
 Best Local Similarity 27.5%; Pred. No. 9.8; Mismatches 26;
 Matches 25; Conservative 16; Indels 2
 Sequence 23, Application US/08525940
 ; Patent No. 5806351
 GENERAL INFORMATION:
 Qy 37 MSATPKHG----LEQCPPAPPVAVGFTGDSGAKETVSDOKR--SOGHNT
 Db 200 ISSTAESGGKPWYFEC--SSLATTTSSGESYDVKLITDLQRQCRTDNHII
 Qy 87 -----HWLTW--VGHLRNHYSSASH 105
 Db 258 AGITAALEANPFLTWRDVHVIVRTSRAGH 288

US-08-525-940-23
 APPLICANT: Wolf, Joseph R.
 TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
 TITLE OF INVENTION: ENCODING SAID PROTEASES
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross & McIntosh
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/525, 940
 FILING DATE: 01-JAN-1995
 CLASSIFICATION: 514
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/7368, 852
 FILING DATE: 01-JAN-1995
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/088, 322
 FILING DATE: 07-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2948-11-C1

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 799 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-525-940-23

RESULT 8
 Best Local Similarity 27.5%; Score 66; DB 2; Length 799;
 Matches 25; Conservative 16; Mismatches 26; Indels 24; Gaps 5;

Qy 37 MSATPKHG---LHQCPPRAPPVATGFGDSAKETVSQDKR--SOQHTWCTLALP-- 86
 Db 221 ISSTAESGKPKPWYLBEC-SSTLAATVYSSGESYDKIITDLRQRCTDNTGTSASAPMA 278

Qy 87 -----HPWLTW-VGHLRNHSSASH 105
 Db 279 AGIALALEANPFLTWRDVQHVIVRTSRAGH 309

RESULT 9
 Best Local Similarity 27.5%; Pred. No. 31;
 Matches 25; Conservative 16; Mismatches 26; Indels 24; Gaps 5;

Qy 37 MSATPKHG---LHQCPPRAPPVATGFGDSAKETVSQDKR--SOQHTWCTLALP-- 86
 Db 221 ISSTAESGKPKPWYLBEC-SSTLAATVYSSGESYDKIITDLRQRCTDNTGTSASAPMA 278

Qy 87 -----HPWLTW-VGHLRNHSSASH 105
 Db 279 AGIALALEANPFLTWRDVQHVIVRTSRAGH 309

RESULT 10
 Best Local Similarity 27.5%; Score 66; DB 2; Length 799;
 Matches 25; Conservative 16; Mismatches 26; Indels 24; Gaps 5;

Qy 37 MSATPKHG---LHQCPPRAPPVATGFGDSAKETVSQDKR--SOQHTWCTLALP-- 86
 Db 221 ISSTAESGKPKPWYLBEC-SSTLAATVYSSGESYDKIITDLRQRCTDNTGTSASAPMA 278

Qy 87 -----HPWLTW-VGHLRNHSSASH 105
 Db 361 AGIALALEANPFLTWRDVQHVIVRTSRAGH 391

RESULT 21
 Sequence 21, Application US/08976838
 Patent No. 5981239
 GENERAL INFORMATION:
 APPLICANT: Franzusoff, Alex
 APPLICANT: Miranda, Luis R.
 APPLICANT: Wolf, Joseph R.
 TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
 TITLE OF INVENTION: ENCODING SAID PROTEASES
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross & McIntosh
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203

COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patientin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/525, 940
 FILING DATE:
 CLASSIFICATION: 514
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/088, 322
 FILING DATE: 07-JUL-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32, 020
 REFERENCE/DOCKET NUMBER: 2848-11-C1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 881 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-976-838-23

Query Match 11.3%; Score 66; DB 2; Length 799;
 Best Local Similarity 27.5%; Pred. No. 31;
 Matches 25; Conservative 16; Mismatches 26; Indels 24; Gaps 5;

Qy 37 MSATPKHG---LHQCPPRAPPVATGFGDSAKETVSQDKR--SOQHTWCTLALP-- 86
 Db 303 ISSTAESGKPKPWYLBEC-SSTLAATVYSSGESYDKIITDLRQRCTDNTGTSASAPMA 360

Qy 87 -----HPWLTW-VGHLRNHSSASH 105
 Db 361 AGIALALEANPFLTWRDVQHVIVRTSRAGH 391

RESULT 21
 Sequence 21, Application US/08976838
 Patent No. 5981239
 GENERAL INFORMATION:
 APPLICANT: Franzusoff, Alex
 TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASE NUCLEIC ACID
 TITLE OF INVENTION: MOLECULES

NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross P.C.
 STREET: 1700 Lincoln St., Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/976,838
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2848-11-C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 881 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-976-838-21

Query Match 11.3%; Score 66; DB 2; Length 881;
 Best Local Similarity 27.5%; Pred. No. 35; Mismatches 26; Indels 24; Gaps 5;
 Matches 25; Conservative 16; N mismatches 26; Indels 24; Gaps 5;
 Patient No. 5886351

GENERAL INFORMATION:
 APPLICANT: Franzusoff, Alex
 APPLICANT: Miranda, Luis R.
 APPLICANT: Wolf, Joseph R.
 TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASES AND GENES
 NUMBER OF SEQUENCES: 25
 TITLE OF INVENTION: ENCODING SAID PROTEASES
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross & McIntosh
 STREET: 1700 Lincoln Street, Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/976,838
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2848-11-C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 915 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-976-838-18

RESULT 11
 US-08-525-940-18
 ; Sequence 18, Application US/08525940

RESULT 12
 US-08-976-838-18
 ; Sequence 18, Application US/08976838
 ; Patent No. 5981259
 GENERAL INFORMATION:
 APPLICANT: Franzusoff, Alex
 TITLE OF INVENTION: CD4+ T-LYMPHOCYTE PROTEASE NUCLEIC ACID
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sheridan Ross P.C.
 STREET: 1700 Lincoln St., Suite 3500
 CITY: Denver
 STATE: Colorado
 COUNTRY: U.S.A.
 ZIP: 80203
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/976,838
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Connell, Gary J.
 REGISTRATION NUMBER: 32,020
 REFERENCE/DOCKET NUMBER: 2848-11-C2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303) 863-9700
 TELEFAX: (303) 863-0223
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 915 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-976-838-18

Query Match 11.3%; Score 66; DB 2; Length 915;
 Best Local Similarity 27.5%; Pred. No. 36; Mismatches 26; Indels 24; Gaps 5;
 Matches 25; Conservative 16; Mismatches 26; Indels 24; Gaps 5;

RESULT 13
 US-09-214-555B-2
 ; Sequence 2, Application US/09214555B
 ; Patent No. 6380171
 ; GENERAL INFORMATION:
 APPLICANT: INSTITUT DE RECHERCHE CLINIQUE DE MONTREAL
 TITLE OF INVENTION: PRO-PROTEIN CONVERTING ENZYME
 FILE REFERENCE: PRO-PROTEIN CONVER ENZ
 CURRENT APPLICATION NUMBER: US/09/214, 555B
 CURRENT FILING DATE: 1999-01-04
 PRIOR APPLICATION NUMBER: 60/021, 008
 PRIOR FILING DATE: 1996-07-26
 PRIORITY NUMBER: 2,203, 745
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2
 LENGTH: 915
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-214-555B-2

Query Match 11.3%; Score 66; DB 4; Length 915;
 Best Local Similarity 27.5%; Pred. No. 36; Mismatches 26; Indels 24; Gaps 5;
 Matches 25; Conservative 16; Mismatches 26; Indels 24; Gaps 5;

QY 37 MSATPKHG---LEQCPPAPPAYGFTGDSGAKETVSQDKR--SQGHTWCTIALP-- 86
 :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 337 ISSTAESKKPPWYLEC--SSTLATRTYSSESBSYDKITIDLRQRCTDNTGTSASAPMA 394
 QY 87 -----HPWLTM--VGLRNHVSASH 105
 :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 395 AGIALALEANPFLWWRDVQHVIVRTSRAGH 425

RESULT 15
 US-08-053-867A-2
 ; Sequence 2, Application US/08053867A
 ; Patent No. 5545545
 ; GENERAL INFORMATION:
 APPLICANT: Gengenbach, Burle G.
 APPLICANT: Somers, David A.
 APPLICANT: Bittel, Douglas C.
 APPLICANT: Shaver, Jonathan M.
 TITLE OF INVENTION: Lysine-Insensitive Maize
 FILE REFERENCE: PRO-PROTEIN CONVERTING ENZYME
 CURRENT APPLICATION NUMBER: US/08/053-867A-1
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSE: Merchant & Gould
 STREET: 3100 NO. 5545545west Center
 CITY: Minneapolis
 STATE: MN
 COUNTRY: USA
 ZIP: 55403
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/053-867A
 FILING DATE: 27-APR-1993
 CLASSIFICATION: 800
 ATTORNEY/AGENT INFORMATION:
 NAME: Woessner, Warren D.
 REGISTRATION NUMBER: 30,440
 REFERENCE/DOCKET NUMBER: 600.263-US-01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 612/332-3300
 TELEFAX: 612/332-3081
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 326 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 IMMEDIATE SOURCE:
 CLONE: Acid Sequence for Maize Mutant DHPS
 ; Patent No. 5545545
 US-08-053-867A-2

Query Match 11.1%; Score 64.5; DB 1; Length 326;
 Best Local Similarity 23.4%; Pred. No. 16; Mismatches 24; Indels 23; Gaps 2;
 Matches 18; Conservative 12; Mismatches 24; Indels 23; Gaps 2;

QY 17 FLIGSHPAIS-----LPLVWTVAGVMSATPKHGLEQCPAPPAP 56
 :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 117 FAVGHNHALHINPYGKISAEGMISHFEAVLPNGPT--IIVNPSRSQAQDIPPEVILAI 173
 QY 57 TGFTGDSGAKETVSQDK 73
 :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
 Db 174 SGYINMAGYRECVRGER 190

Query Match 11.3%; Score 66; DB 4; Length 915;
 Best Local Similarity 27.5%; Pred. No. 36; Mismatches 26; Indels 24; Gaps 5;
 Matches 25; Conservative 16; Mismatches 26; Indels 24; Gaps 5;

Search completed: June 17, 2003, 10:38:13
 Job time : 28 secs

Om protein - protein search, using sw model

Run on: June 17, 2003, 10:37:01 ; Search time 47 Seconds
 (without alignments)
 238.854 Million cell updates/sec

Title: US-09-955-807-2
Perfect score: 582
Sequence: 1 MGYSPEMPCAHPIGLFLIG.....PHFWPLTWVGHLRNHVSSASH 105

Scoring table: BLOSUM62
Gapext 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pupbaa/US08_NEW_PUB.pep:*
 2: /cgn2_6/ptodata/1/pupbaa/PC1_NEW_PUB.pep:*
 3: /cgn2_6/ptodata/1/pupbaa/US06_NEW_PUB.pep:*
 4: /cgn2_6/ptodata/1/pupbaa/US06_PUBCOMB.pep:*
 5: /cgn2_6/ptodata/1/pupbaa/US07_NEW_PUB.pep:*
 6: /cgn2_6/ptodata/1/pupbaa/US07_PUBCOMB.pep:*
 7: /cgn2_6/ptodata/1/pupbaa/PCTUS_PUBCOMB.pep:*
 8: /cgn2_6/ptodata/1/pupbaa/US08_PUBCOMB.pep:*
 9: /cgn2_6/ptodata/1/pupbaa/US09_NEW_PUB.pep:*
 10: /cgn2_6/ptodata/1/pupbaa/US09_PUBCOMB.pep:*
 11: /cgn2_6/ptodata/1/pupbaa/US10_NEW_PUB.pep:*
 12: /cgn2_6/ptodata/1/pupbaa/US10_PUBCOMB.pep:*
 13: /cgn2_6/ptodata/1/pupbaa/US60_NEW_PUB.pep:*
 14: /cgn2_6/ptodata/1/pupbaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	582	100.0	105	10 US-09-955-807-2
2	582	100.0	105	10 US-09-955-807-14
3	443	76.1	80	10 US-09-955-807-16
4	443	76.1	89	10 US-09-955-807-15
5	439	75.4	79	10 US-09-955-807-3
6	428	73.5	77	10 US-09-955-807-4
7	375	64.4	65	10 US-09-955-807-5
8	372	63.9	65	10 US-09-955-807-10
9	246	42.3	43	10 US-09-955-807-12
10	229	39.3	42	10 US-09-955-807-9
11	114	19.5	20	10 US-09-955-807-8
12	103	17.7	20	10 US-09-955-807-11
13	79	13.6	9	10 US-09-955-807-1
14	76	13.1	380	10 US-09-955-807-14
15	70	12.0	76	9 US-09-955-807-14
16	70	12.0	76	9 US-10-040-862-876
17	70	12.0	1616	9 US-09-712-363-262
18	69.5	11.9	147	9 US-10-016-63A-115
19	68	11.7	73	9 US-10-040-862-1130

ALIGNMENTS

RESULT 1

US-09-955-807-2

; Sequence 2, Application US/09955807

; Patent No. US20020132996A1

; GENERAL INFORMATION:

; APPLICANT: Lok, S.I.

; APPLICANT: Steppard, Paul O.

; APPLICANT: Kindvogel, Wayne

; APPLICANT: Bort, Susan J.

; TITLE OF INVENTION: Secretory Protein-48

; FILE REFERENCE: 98-17C1

; CURRENT APPLICATION NUMBER: US/09/955, 807

; CURRENT FILING DATE: 2001-09-19

; PRIOR APPLICATION NUMBER: 60/102, 679

; PRIOR FILING DATE: 1998-10-01

; PRIOR APPLICATION NUMBER: 09/410, 603

; PRIOR FILING DATE: 1999-10-01

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: Fast-SEQ for Windows Version 3.0

; SEQ ID NO: 2

; LENGTH: 105

; TYPE: PRT

; ORGANISM: Homo sapiens

Query Match Score 582; DB 10; Length 105;

Best local Similarity 100.0%; Pred. No. 1.5e-51;

Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYSPEMPCAHPIGLFLIGLHPALSLPLVWVAGVMSATPKIGLEQCPAPPAPPAVFTGFT 60

Db 1 MGYSPEMPCAHPIGLFLIGLHPALSLPLVWVAGVMSATPKIGLEQCPAPPAPPAVFTGFT 60

QY 61 GDSGAKETVSQDKRSQGHWTCLALPHWLTVWGHLRNHVSSASH 105

Db 61 GDSGAKETVSQDKRSQGHWTCLALPHWLTVWGHLRNHVSSASH 105

RESULT 2

US-09-955-807-14

; Sequence 14, Application US/09955807

; Patent No. US20020132996A1

RESULT 4

RESULT 6
US-09-955-807-4
; Sequence 4, Application US/09955807
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
FILE REFERENCE: 98-17C1
CURRENT APPLICATION NUMBER: US/09/955, 807
CURRENT FILING DATE: 2001-09-19
PRIORITY NUMBER: 60/102, 679
PRIORITY NUMBER: 1998-10-01
PRIORITY APPLICATION NUMBER: 09/410, 603
PRIORITY FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 4
LENGTH: 77
TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-807-4

Query Match 73.5%; Score 428; DB 10; Length 77;
Best Local Similarity 100.0%; Pred. No. 3.5e-36;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LAVTVAGVMSATPKHGLEQCPAPPAPAVTGFIDGDSGAKETVQDKRSQHTWCTLALPHP 88
Db 1 LAVTVAGVMSATPKHGLEQCPAPPAPAVTGFIDGDSGAKETVQDKRSQHTWCTLALPHP 60

QY 89 WLTWVGHLRNHYSSASH 105
Db 61 WLTWVGHLRNHYSSASH 77

RESULT 7
US-09-955-807-5
Sequence 5, Application US/09955807
Patent No. US20020132996A1
GENERAL INFORMATION:
APPLICANT: Lok, Si
APPLICANT: Sheppard, Paul O.
APPLICANT: Kindsvogel, Wayne
APPLICANT: Bort, Susan J.
TITLE OF INVENTION: Secretory Protein-48
FILE REFERENCE: 98-17C1
CURRENT APPLICATION NUMBER: US/09/955, 807
CURRENT FILING DATE: 2001-09-19
PRIORITY NUMBER: 60/102, 679
PRIORITY FILING DATE: 1998-10-01
PRIORITY APPLICATION NUMBER: 09/410, 603
PRIORITY FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 5
LENGTH: 65
TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-807-5

Query Match 64.4%; Score 375; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 6.4e-31;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 PKHGLEQCPAPPAPAVTGFIDGDSGAKETVQDKRSQHTWCTLALPHWLTWVGHLRHV 100
Db 1 PKHGLEQCPAPPAPAVTGFIDGDSGAKETVQDKRSQHTWCTLALPHWLTWVGHLRHV 60

RESULT 8
US-09-955-807-10
Sequence 10, Application US/09955807
Patent No. US20020132996A1
GENERAL INFORMATION:
APPLICANT: Lok, Si
APPLICANT: Sheppard, Paul O.
APPLICANT: Kindsvogel, Wayne
APPLICANT: Bort, Susan J.
TITLE OF INVENTION: Secretory Protein-48
FILE REFERENCE: 98-17C1
CURRENT APPLICATION NUMBER: US/09/955, 807
CURRENT FILING DATE: 2001-09-19
PRIORITY NUMBER: 60/102, 679
PRIORITY APPLICATION NUMBER: 1998-10-01
PRIORITY FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 10
LENGTH: 65
TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-807-10

Query Match 63.9%; Score 372; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.3e-30;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 SATPKHGLEQCPAPPAPAVTGFIDGDSGAKETVQDKRSQHTWCTLALPHWLTWVGHLR 97
Db 1 SATPKHGLEQCPAPPAPAVTGFIDGDSGAKETVQDKRSQHTWCTLALPHWLTWVGHLR 60

QY 98 NHVSS 102
Db 61 NHVSS 65

RESULT 9
US-09-955-807-12
Sequence 12, Application US/09955807
Patent No. US20020132996A1
GENERAL INFORMATION:
APPLICANT: Lok, Si
APPLICANT: Sheppard, Paul O.
APPLICANT: Kindsvogel, Wayne
APPLICANT: Bort, Susan J.
TITLE OF INVENTION: Secretory Protein-48
FILE REFERENCE: 98-17C1
CURRENT APPLICATION NUMBER: US/09/955, 807
CURRENT FILING DATE: 2001-09-19
PRIORITY NUMBER: 60/102, 679
PRIORITY FILING DATE: 1998-10-01
PRIORITY APPLICATION NUMBER: 09/410, 603
PRIORITY FILING DATE: 1999-10-01
NUMBER OF SEQ ID NOS: 17
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 12
LENGTH: 43
TYPE: PRT
ORGANISM: Homo sapiens
US-09-955-807-12

Query Match 42.3%; Score 246; DB 10; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.1e-18;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 TGDSGAKETVQDKRSQHTWCTLALPHWLTWVGHLRHV 102

```

Db          |||||||TGDSAKETVSQDKRSQGHTWCTALPHPHWTWGHLRNHYSS 43
Db          ; Sequence 11, Application US/09955807
Db          ; Patent No. US20020132996A1
Db          ; GENERAL INFORMATION:
Db          ; APPLICANT: Lok, Si
Db          ; APPLICANT: Sheppard, Paul O.
Db          ; APPLICANT: Kindsvogel, Wayne
Db          ; APPLICANT: Bort, Susan J.
Db          ; TITLE OF INVENTION: Secretory Protein-48
Db          ; FILE REFERENCE: 98-17C1
Db          ; CURRENT APPLICATION NUMBER: US/09/955, 807
Db          ; CURRENT FILING DATE: 2001-09-19
Db          ; PRIOR APPLICATION NUMBER: 60/102, 679
Db          ; PRIOR FILING DATE: 1998-10-01
Db          ; PRIOR APPLICATION NUMBER: 09/410, 603
Db          ; PRIOR FILING DATE: 1999-10-01
Db          ; NUMBER OF SEQ ID NOS: 17
Db          ; SOFTWARE: FastSEQ for Windows Version 3.0
Db          ; SEQ ID NO: 9
Db          ; LENGTH: 42
Db          ; ORGANISM: Homo sapiens
Db          ; US-09-955-807-9

Query Match   39, 3%; Score 229; DB 10; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.1e-16; Mismatches 0; Indels 0; Gaps 0;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          38 SATPKHGLEOCPPAPPATVGTGFTGDSAKETVSQDKRSQHT 79
Qy          |||||TGDSAKETVSQDKRSQHT 42
Db          1 SATPKHGLEOCPPAPPATVGTGFTGDSAKETVSQDKRSQHT 42

RESULT 11
US-09-955-807-9
; Sequence 8, Application US/09955807
; Patent No. US20020132996A1
; GENERAL INFORMATION:
; APPLICANT: Lok, Si
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Bort, Susan J.
; TITLE OF INVENTION: Secretory Protein-48
; FILE REFERENCE: 98-17C1
; CURRENT APPLICATION NUMBER: US/09/955, 807
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/102, 679
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 09/110, 603
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 9
; LENGTH: 42
; ORGANISM: Homo sapiens
; US-09-955-807-9

Query Match   17, 7%; Score 103; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00044; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          60 TGDSAKETVSQDKRSQHT 79
Qy          |||||TGDSAKETVSQDKRSQHT 20
Db          1 TGDSAKETVSQDKRSQHT 20

RESULT 13
US-09-764-891-4761
; Sequence 4761, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764, 891
; CURRENT FILING DATE: 2001-01-17
; PRIOR application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10331
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 4761
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (25)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (29)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (42)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-891-4761

Query Match   13, 6%; Score 79; DB 9; Length 136;
Best Local Similarity 28.4%; Pred. No. 1.1e-05; Mismatches 5; Indels 42; Gaps 5;
Matches 33; Conservative 5; Mismatches 36; Indels 42; Gaps 5;

Qy          7 PMPCKHGLEOCPPAPPATVGTGFTGDSAKETVSQDKRSQHT 61
Qy          |||||TGDSAKETVSQDKRSQHT 61
Db          33 PAPC-----LGTVHXDQ---VTPGCTFSADPSPWPQGNHNGPAPPAS--- 78
Db          1 SATPKHGLEOCPPAPPATVGTGFTGDSAKETVSQDKRSQHT 20

RESULT 12

```


THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model
 Copyright (c) 1993 - 2003 Compugen Ltd.
 GenCore version 5.1.6

Run on: June 17, 2003, 10:25:05 ; search time '0 seconds
 (without alignments)
 199.876 Million cell updates/second

Title: US-09-955-807-2
perfect score: 582
Sequence: MGYSSPPMPCARPLSLFLLG.....PHPWLAQWVGLRHNHYSSASH 105

Scoring table: BLOSSUM62
 Gapext 10.0 , gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database :

```

1: /SIDS2/gcadata/geneseq/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcadata/geneseq/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcadata/geneseq/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcadata/geneseq/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcadata/geneseq/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcadata/geneseq/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcadata/geneseq/geneseqp-emb1/AA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	582	100.0	105	21 AAY79436	Human secretary
2	443	76.1	80	21 AAY79440	Human Secretary
3	443	76.1	89	21 AAY79446	Human Secretary
4	439	75.4	79	21 AAY79437	Human secretary
5	428	73.5	77	21 AAY79438	Human Secretary
6	375	64.9	65	21 AAY79439	Human secretary
7	372	63.9	65	21 AAY79443	Human secretary
8	245	42.3	43	21 AAY79445	Human Secretary
9	229	39.3	42	21 AAY79442	Human secretary
10	114	19.6	20	21 AAY79441	Human secretary

11	103	17.7	20	21	AAV79444	Human secretory pr
12	89	15.3	21	22	ABB64617	Drosophila melanog
13	79.5	13.7	21	22	ABR42156	Human ORFX ORF1920
14	79.5	13.7	23	23	ABR97893	Human secretory Po
15	79	13.6	22	22	AMW5103	Human reproductive
16	79	13.6	188	21	ABR38275	Human secreted pro
17	79	13.6	188	21	ABR38277	Gene 19 human secr
18	77	13.2	21	21	ABR41578	Human ORFX ORF1442
19	75.5	13.0	75	22	ABX49832	Propionibacterium
20	73.5	12.6	138	20	AVX37090	Amino acid sequenc
21	73.5	12.6	941	23	ABX38686	Cricetulus griseus
22	72.5	12.5	364	21	ABX53186	Macaca mulatta rha
23	72	12.4	141	22	ABX79666	Human protein SBQ
24	72	12.4	572	22	ABX32685	Novel human secret
25	71.5	12.3	146	22	ABX61349	Propionibacterium
26	71.5	12.3	989	22	ABX90769	Human shear stress
27	71	12.2	93	22	ABX19572	Novel human diagno
28	71	12.2	415	22	ABG09527	Novel human diagno
29	71	12.2	900	21	ABX42321	Human ORFX ORF2085
30	71	12.2	1217	22	ABX09876	Novel human diagno
31	71	12.2	1239	22	ABG09877	Novel human diagno
32	71	12.2	1964	20	ABX95557	Mus musculus notch
33	70.5	12.1	164	21	ABX44836	Human Perlephin-AR
34	70.5	12.1	813	22	ABG25423	Novel human diagno
35	70	12.0	76	22	ABM80512	Human haematologic
36	70	12.0	1615	22	ABG81211	Mycobacterium tube
37	69.5	11.9	186	22	AAM24399	Human EST encoded
38	69.5	11.9	305	22	ABU58771	Propionibacterium
39	69	11.9	361	22	ABG10502	Novel human diagno
40	69	11.9	423	20	AAY03228	Amino acid sequenc
41	69	11.9	423	21	ABR10280	Murine adult thymu
42	69	11.9	447	22	ABR66646	Drosophila melanog
43	68.5	11.8	137	23	ABR89082	Human polypeptide
44	68.5	11.8	275	22	AAY42298	Propionibacterium
45	68.5	11.8	301	22	ABM42100	Human polypeptide

FT /note= "Zs1g48 mature polypeptide"
 XX WO200018796-A2.
 PN 01-OCT-1999; 99WO-US22970.
 XX PR 01-OCT-1998; 98US-0164740.
 PD 06-APR-2000.
 XX PA (ZYMO) ZYMOGENETICS INC.
 XX PI Lok S, Sheppard PO;
 XX DR WPI; 2000-303441/26.
 XX PS Claim 1; Page 138; 141pp; English.
 XX PT Human secretory protein 48 polypeptides and polynucleotides useful for promoting leukocyte proliferation and for treating immunosuppressed individuals -
 XX PT individuals -
 XX Claim 1; Page 138; 141pp; English.
 The present sequence is that of a human secretory protein 48 (zs1g48) polypeptide. The invention provides zs1g48 polypeptides, antibodies, and anti-idiotypic antibodies. Zs1g48 stimulates the proliferation of peripheral blood mononuclear cells, i.e. T-cells, B-cells and monocytes. The invention provides zs1g48 polynucleotides, polypeptides, antibodies and anti-idiotypic antibodies. Zs1g48 stimulates the proliferation of leukocytes in both a mixed leukocyte reaction and in an unmixed leukocyte reaction. It can be used to promote proliferation of peripheral blood leukocytes and to treat low leukocyte counts in individuals (claimed). This can be useful in treating cancer patients whose leukocytes have been depleted by chemotherapy, radiation or receiving bone marrow transplants to promote proliferation of leukocytes produced by the transplanted marrow. It could also be useful in treating immunosuppressed individuals such as the elderly or HIV infected individuals, or used as a vaccine adjuvant.
 CC marrow transplants to promote proliferation of leukocytes produced by the transplanted marrow. It could also be useful in treating immunosuppressed individuals such as the elderly or HIV infected individuals, or used as a vaccine adjuvant.
 CC individuals, or used as a vaccine adjuvant.
 XX Sequence 89 AA;
 Query Match 76.1%; Score 443; DB 21; Length 89;
 Best Local Similarity 100.0%; Pred. No. 1.1e-39; Mismatches 0; Indels 0; Gaps 0;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Qy 26 SLPVLTWVAGMSATPKHGLEQCPAPPVATGFGDSGAKETVSQDKRSQHWTALP 85
 Db 10 SLPVLTWVAGMSATPKHGLEQCPAPPVATGFGDSGAKETVSQDKRSQHWTALP 69
 SQ Sequence 79 AA;
 Query Match 75.4%; Score 439; DB 21; Length 79;
 Best Local Similarity 100.0%; Pred. No. 2.5e-39; Mismatches 0; Indels 0; Gaps 0;
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Qy 27 LPLVWVAGMSATPKHGLEQCPAPPVATGFGDSGAKETVSQDKRSQHWTALP 86
 Db 1 LPLVWVAGMSATPKHGLEQCPAPPVATGFGDSGAKETVSQDKRSQHWTALP 60
 SQ Sequence 79 AA;
 Query Match 75.4%; Score 439; DB 21; Length 79;
 Best Local Similarity 100.0%; Pred. No. 2.5e-39; Mismatches 0; Indels 0; Gaps 0;
 Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Qy 86 PHPWLITWVGHLRNHSSASH 105
 Db 70 PHPWLITWVGHLRNHSSASH 89
 SQ Sequence 77 AA;
 RESULT 4
 AAY79437
 ID AAY79437 standard; Protein; 79 AA.
 XX
 AC AAY79437;
 XX
 DT 01-AUG-2000 (first entry)
 XX Human secretory protein 48 (zs1g48) mature polypeptide.
 DE Human secretory protein 48 (zs1g48) mature polypeptide.
 KW Secretory protein 48; Zs1g48; human; leukocyte; proliferation; immunostimulant; adjuvant; therapy.
 KW Homo sapiens.
 OS Homo sapiens.
 PN WO200018796-A2.
 PR 01-OCT-1999; 99WO-US22970.
 XX PR 01-OCT-1998; 98US-0164740.
 XX PA (ZYMO) ZYMOGENETICS INC.
 XX
 FT 01-OCT-1999; 99WO-US22970.
 XX PR 01-OCT-1998; 98US-0164740.
 XX PA (ZYMO) ZYMOGENETICS INC.
 XX PI Lok S, Sheppard PO;
 XX DR WPI; 2000-303441/26.
 XX PS Claim 1; Page 129; 141pp; English.
 XX PT Human secretory protein 48 polypeptides and polynucleotides useful for promoting leukocyte proliferation and for treating immunosuppressed individuals -
 XX PT individuals -
 XX PS Claim 1; Page 129; 141pp; English.
 The present sequence is that of human secretory protein 48 (zs1g48) mature polypeptide, i.e. lacking the signal sequence (see AAY79437). Zs1g48 mature protein stimulates the proliferation of peripheral blood mononuclear cells i.e. T-cells, B-cells and monocytes. The invention provides zs1g48 polynucleotides, polypeptides, antibodies and anti-idiotypic antibodies. Zs1g48 stimulates the proliferation of leukocytes in both a mixed leukocyte reaction and in an unmixed leukocyte reaction. It can be used to promote proliferation of peripheral blood leukocytes and to treat low leukocyte counts in individuals (claimed). This can be useful in treating cancer patients whose leukocytes have been depleted by chemotherapy, radiation or receiving bone marrow transplants to promote proliferation of leukocytes produced by the transplanted marrow. It could also be useful in treating immunosuppressed individuals such as the elderly or HIV infected individuals, or used as a vaccine adjuvant.

CC stimulates the proliferation of peripheral blood mononuclear cells, i.e. T-cells, B-cells and monocytes. The invention provides Zsig48 polynucleotides, polypeptides, antibodies and anti-idiotypic antibodies. Zsig48 stimulates the proliferation of leukocytes in both a mixed leukocyte reaction and in an unmixed leukocyte reaction. It can be used to promote proliferation of peripheral blood leukocytes and to treat low leukocyte counts in individuals (claimed). This can be useful in treating cancer patients whose leukocytes have been depleted by chemotherapy, radiation or illness. Zsig48 can be administered to patients receiving bone marrow transplants to promote proliferation of leukocytes produced by the transplanted marrow. It could also be useful in treating immunosuppressed individuals such as the elderly or HIV infected individuals, or used as a vaccine adjuvant. Anti-Zsig48 antibodies such as the present sequence and used to detect Zsig48 polypeptides.

XX Sequence 65 AA:

Query Match 63.9%; Score 372; DB 21; Length 65;

Best Local Similarity 100.0%; Pred. No. 2.7e-32; Mismatches 0; Indels 0; Gaps 0;

QY 38 SATPKHLEQCPAPPAVTGTFGDSAKETVSQDRRSQGHWTWGLHLR 97

1 SATPKHLEQCPAPPAVTGTFGDSAKETVSQDRRSQGHWTWGLHLR 60

OY 98 NIVSS 102

Db 61 NIVSS 65

RESULT 8

AAV7945 AAV7945 standard; Peptide; 43 AA.

AC AAV7945; AAV7945;

DE 01-AUG-2000 (first entry)

DR Human secretory protein 48 (Zsig48) antigenic peptide.

KW Secretory protein 48; Zsig48; human; leukocyte; proliferation; immunostimulant; adjuvant; therapy; antigen; immunogen; antibody.

OS Homo sapiens.

XX WO200118796-A2.

PD 06-APR-2000.
XX
PF 01-OCT-1999; 99WO-US22970.

PR 01-OCT-1998; 98US-0164740.

PA (ZYMO) ZYMOGENETICS INC.
XX
PI Lok S, Sheppard PO;
XX
DR WPI; 2000-303441-26.

PT Human secretory protein 48 polypeptides and polynucleotides useful for promoting leukocyte proliferation and for treating immunosuppressed individuals -
XX
PS Claim 2; Page 131; 141pp; English.

XX Human secretory protein 48 polypeptides and polynucleotides useful for promoting leukocyte proliferation and for treating immunosuppressed individuals -
XX
PS Claim 2; Page 132; 141pp; English.

CC the present sequence is that of an antigenic peptide of human secretory protein 48 (Zsig48, see also AAV79436-40), a protein that promotes the proliferation of peripheral blood mononuclear cells, i.e. T-cells, B-cells and monocytes. The invention provides Zsig48 polynucleotides, polypeptides, antibodies and anti-idiotypic antibodies. Zsig48 stimulates the proliferation of leukocytes in both a mixed leukocyte reaction and in an unmixed leukocyte reaction. It can be used to promote proliferation of peripheral blood leukocytes and to treat low leukocyte counts in individuals (claimed). This can be useful in treating cancer patients whose leukocytes have been depleted by chemotherapy, radiation or illness. Zsig48 can be administered to patients receiving bone marrow transplants to promote proliferation of leukocytes produced by the transplanted marrow. It could also be useful in treating immunosuppressed individuals such as the elderly or HIV infected individuals, or used as a vaccine adjuvant. Anti-Zsig48 antibodies such as the present sequence and used to detect Zsig48 polypeptides.

XX Sequence 43 AA:

CC both a mixed leukocyte reaction and in an unmixed leukocyte reaction. It can be used to promote proliferation of peripheral blood leukocytes and to treat low leukocyte counts in individuals (claimed). This can be useful in treating cancer patients whose leukocytes have been depleted by chemotherapy, radiation or illness. Zsig48 can be administered to patients receiving bone marrow transplants to promote proliferation of leukocytes produced by the transplanted marrow. It could also be useful in treating immunosuppressed individuals such as the elderly or HIV infected individuals, or used as a vaccine adjuvant. Anti-Zsig48 antibodies such as the present sequence and used to detect Zsig48 polypeptides.

XX Sequence 43 AA:

Query Match 42.3%; Score 246; DB 21; Length 43;

Best Local Similarity 100.0%; Pred. No. 4.4e-19; Mismatches 0; Indels 0; Gaps 0;

QY 60 TGDSGAKETVSQDRRSQGHWTWGLHLRNHYSS 102

1 TGDSGAKETVSQDRRSQGHWTWGLHLRNHYSS 43

RESULT 9

AAV7942 AAV7942 standard; Peptide; 42 AA..

AC AAV7942;

DE 01-AUG-2000 (first entry)

DR Human secretory protein 48 (Zsig48) antigenic peptide.

XX Secretory protein 48; Zsig48; human; leukocyte; proliferation; immunostimulant; adjuvant; therapy; antigen; immunogen; antibody.

OS Homo sapiens.

XX WO200118796-A2.

PD 06-APR-2000.

XX
PF 01-OCT-1999; 99WO-US22970.

XX
PR 01-OCT-1998; 98US-0164740.

XX
PA (ZYMO) ZYMOGENETICS INC.

XX
PI Lok S, Sheppard PO;

XX
DR WPI; 2000-303441-26.

PT Human secretory protein 48 polypeptides and polynucleotides useful for promoting leukocyte proliferation and for treating immunosuppressed individuals -
XX
PS Claim 2; Page 131; 141pp; English.

CC The present sequence is that of an antigenic peptide of human secretory protein 48 (Zsig48, see also AAV79436-40), a protein that stimulates the proliferation of peripheral blood mononuclear cells, i.e. T-cells, B-cells and monocytes. The invention provides Zsig48 polynucleotides, polypeptides, antibodies and anti-idiotypic antibodies. Zsig48 stimulates the proliferation of leukocytes in both a mixed leukocyte reaction and in an unmixed leukocyte reaction. It can be used to promote proliferation of peripheral blood leukocytes and to treat low leukocyte counts in individuals (claimed). This can be useful in treating cancer patients whose leukocytes have been depleted by chemotherapy, radiation or illness. Zsig48 can be administered to patients receiving bone marrow transplants to promote proliferation of leukocytes produced by the transplanted marrow. It could also be useful in treating immunosuppressed individuals such as the elderly or HIV infected individuals, or used as a vaccine adjuvant. Anti-Zsig48 antibodies such as the present sequence and used to detect Zsig48 polypeptides.

CC immunosuppressed individuals such as the elderly or HIV infected individuals, or used as a vaccine adjvant. Anti-zs1g48 antibodies can be produced using antigenic Zs1g48 epitope-bearing peptides such as the present sequence and used to detect Zs1g48 polypeptides.

XX Sequence 42 AA;

Query Match 39.3%; Score 229; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 38 SATPRHGLEQCPAPPAPAVTGFIDSGAKETVSQDKRSQHT 79
1 SATPRHGLEQCPAPPAPAVTGFIDSGAKETVSQDKRSQHT 42

SQ RESULT 10

AAY79441 standard; Peptide; 20 AA.

XX AAY79441;

XX DT 01-AUG-2000 (first entry)

XX Human secretory protein 48 (zs1g48) antigenic peptide.

XX Secretory protein 48; Zs1g48; human; leukocyte; proliferation;

KW immunostimulant; adjuvant; therapy; antigen; immunogen; antibody.

XX OS Homo sapiens.

XX WO200018796-A2.

XX PN 06-APR-2000.

XX PD 01-OCT-1999; 99WO-US22970.

XX PF (ZYMO) ZYMOGENETICS INC.

XX PT Lok S, Sheppard PO;

XX DR WPI; 2000-303441/26.

XX XX Human secretory protein 48 polypeptides and polynucleotides useful for

PT promoting leukocyte proliferation and for treating immunosuppressed

PT individuals -

XX PS Claim 2; Page 132; 141pp; English.

XX CC The present sequence is that of an antigenic peptide of human

CC secretory protein 48 (Zs1g48, see also AAY79436-40), a protein that

CC stimulates the proliferation of peripheral blood mononuclear cells,

CC i.e., T-cells, B-cells and monocytes. The invention provides Zs1g48

CC polynucleotides, polypeptides, antibodies and anti-idiotypic

CC antibodies. Zs1g48 stimulates the proliferation of leukocytes in

CC both a mixed leukocyte reaction and in an unmixed leukocyte

CC reaction. It can be used to promote proliferation of peripheral

CC blood leukocytes and to treat low leukocyte counts in individuals

CC (claimed). This can be useful in treating cancer patients whose

CC leukocytes have been depleted by chemotherapy, radiation or illness.

CC Zs1g48 can be administered to patients receiving bone marrow

CC transplants to promote proliferation of leukocytes produced by the

CC transplanted marrow. It could also be useful in treating

CC immunosuppressed individuals such as the elderly or HIV infected

CC individuals, or used as a vaccine adjvant. Anti-Zs1g48 antibodies

CC can be produced using antigenic Zs1g48 epitope-bearing peptides

CC such as the present sequence and used to detect Zs1g48 polypeptides.

CC XX Sequence 20 AA;

CC SQ Sequence 20 AA;

Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 38 SATPRHGLEQCPAPPAPAVTGFIDSGAKETVSQDKRSQHT 57
1 SATPRHGLEQCPAPPAPAVT

Db 1 SATPRHGLEQCPAPPAPAVT

RESULT 11

AAY79444 standard; Peptide; 20 AA.

XX AAY79444;

XX DT 01-AUG-2000 (first entry)

XX DE Human secretory protein 48 (zs1g48) antigenic peptide.

XX KW Secretory Protein 48; Zs1g48; human; leukocyte; proliferation;

KW immunostimulant; adjuvant; therapy; antigen; immunogen; antibody.

XX OS Homo sapiens.

XX PN WO200018796-A2.

XX PD 06-APR-2000.

XX PR 01-OCT-1999; 99WO-US22970.

XX PI Lok S, Sheppard PO;

XX DR DR WPI; 2000-303441/26.

XX XX Human secretory protein 48 polypeptides and polynucleotides useful for

PT promoting leukocyte proliferation and for treating immunosuppressed

PT individuals -

XX PS Claim 2; Page 132; 141pp; English.

XX CC The present sequence is that of an antigenic peptide of human

CC secretory protein 48 (Zs1g48, see also AAY79436-40), a protein that

CC stimulates the proliferation of peripheral blood mononuclear cells,

CC i.e., T-cells, B-cells and monocytes. The invention provides Zs1g48

CC polynucleotides, polypeptides, antibodies and anti-idiotypic

CC antibodies. Zs1g48 stimulates the proliferation of leukocytes in

CC both a mixed leukocyte reaction and in an unmixed leukocyte

CC reaction. It can be used to promote proliferation of peripheral

CC blood leukocytes and to treat low leukocyte counts in individuals

CC (claimed). This can be useful in treating cancer patients whose

CC leukocytes have been depleted by chemotherapy, radiation or illness.

CC Zs1g48 can be administered to patients receiving bone marrow

CC transplants to promote proliferation of leukocytes produced by the

CC transplanted marrow. It could also be useful in treating

CC immunosuppressed individuals such as the elderly or HIV infected

CC individuals, or used as a vaccine adjvant. Anti-Zs1g48 antibodies

CC can be produced using antigenic Zs1g48 epitope-bearing peptides

CC such as the present sequence and used to detect Zs1g48 polypeptides.

CC XX Sequence 20 AA;

CC SQ Sequence 20 AA;

Query Match 39.3%; Score 229; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 38 SATPRHGLEQCPAPPAPAVTGFIDSGAKETVSQDKRSQHT 79
1 SATPRHGLEQCPAPPAPAVT

Db 1 SATPRHGLEQCPAPPAPAVT

Query Match 19.6%; Score 114; DB 21; Length 20;

Qy	62 DSGAKETVSDQDKRSQGH	78	DR	N-PSDB; ABL99890.
	: _ _ _ _ _ _		XX	
Db	ABB97893 standard; Protein; 598 AA.		PT	Polynucleotide sequences encoding human secretory proteins useful for
	XX		PT	gene therapy of e.g. genetic deficiency disorders, cancers, and
	ABB97893;		PT	diseases caused by intracellular parasites -
XX			XX	
RESULT 14			PS	Claim 29; Page 459-460; 585pp; English.
ABB97893			XX	
ID	ABB97893		CC	The invention comprises the amino acid and coding sequences of human
XX			CC	secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are
AC			CC	useful for treating a disease or condition associated with the expression
XX			CC	of functional SPTM. The SPTM DNA sequences are useful for somatic or
DT	03-OCT-2002 (first entry)		CC	germline gene therapy to correct a genetic deficiency (e.g. severe
DE	Human secretory polypeptide (SPTM) 145.		CC	combined immunodeficiency). The SPTM DNA sequences are also useful in
XX			CC	providing protection against intracellular parasites (e.g. fungal
KW	Human; secretory protein; secretory polynucleotides; SPTM;		CC	parasites and protozoan parasites). The SPTM DNA sequences are useful for
KW	SPTM-related disease; somatic gene therapy; germline gene therapy;		CC	functional SPTM. The SPTM DNA sequences are also useful for diagnosing cell
KW	severe combined immunodeficiency; intracellular parasite protection;		CC	immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's
KW	fungal parasite; protozoan parasite; cell proliferative disorder; cancer;		CC	disease), motor neuron disorders, demyelinating diseases (e.g. multiple
KW	immune disorder; AIDS; neurological disorder; Parkinson's disease;		CC	sclerosis), meningitis, abscesses, prion diseases, cerebral palsy,
KW	motor neuron disorder; demyelinating disease; multiple sclerosis;		CC	neurokeletal disorders, peripheral nervous system disorders,
KW	meningitis; abscess; prion diseases; cerebral palsy;		CC	dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental
KW	neuroskeletal disorder; peripheral nervous system disorder;		CC	disorders (e.g. Tourette's syndrome). Amino acid sequences ABB9749 -
KW	dermatomyositis; polymyositis; myopathy; myasthenia gravis;		CC	ABB9793 represent human secretory proteins of the invention.
KW	mental disorder; Tourette's syndrome.		XX	
XX			Sequence	598 AA;
OS	Homo sapiens.			
XX				
PN	WO200220756-A2.			
XX				
PD	14-MAR-2002.			
XX				
PF	30-AUG-2001; 2001WO-US27297.			
XX				
PR	05-SEP-2000; 2000US-223747P.			
PR	05-SEP-2000; 2000US-229748P.			
PR	05-SEP-2000; 2000US-229750P.			
PR	05-SEP-2000; 2000US-229751P.			
PR	05-SEP-2000; 2000US-230016P.			
PR	05-SEP-2000; 2000US-230583P.			
PR	06-SEP-2000; 2000US-230503P.			
PR	06-SEP-2000; 2000US-230514P.			
PR	06-SEP-2000; 2000US-230515P.			
PR	06-SEP-2000; 2000US-230517P.			
PR	06-SEP-2000; 2000US-230518P.			
PR	06-SEP-2000; 2000US-230519P.			
PR	06-SEP-2000; 2000US-230595P.			
PR	06-SEP-2000; 2000US-230596P.			
PR	06-SEP-2000; 2000US-230597P.			
PR	06-SEP-2000; 2000US-230599P.			
PR	06-SEP-2000; 2000US-230610P.			
PR	06-SEP-2000; 2000US-230864P.			
PR	06-SEP-2000; 2000US-230865P.			
PR	06-SEP-2000; 2000US-230988P.			
PR	06-SEP-2000; 2000US-230989P.			
PR	06-SEP-2000; 2000US-230990P.			
PR	07-SEP-2000; 2000US-230896P.			
PR	07-SEP-2000; 2000US-230897P.			
PR	07-SEP-2000; 2000US-230951P.			
PR	07-SEP-2000; 2000US-231163P.			
PR	07-SEP-2000; 2000US-231832P.			
XX				
PA	(INCY-) INCYTE GENOMICS INC.			
XX				
PI	Stuart J, Lincoln SE, Altis CM, DuFour GE, Chalup MS, Hillman JL,			
PI	Jones AL, Yu JY, Wright RJ, Gletzner D, Liu TF, Yap PE, Dahl CR;			
PI	Moniyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;			
PI	Gerstlin EH, Peratta CH, David MH, Panzer SR, Flores V, Darfo A;			
PI	Marwaha R, Chen AJ, Chang SC, Au AP, Inman RR;			
XX				
DR	07-JUN-2000; 2000US-0215135.			
PR	07-JUL-2000; 2000US-0216647.			
PR	WPI: 2002-315658/35.			

PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0227182.
 PR 14-AUG-2000; 2000US-0227578.
 PR 18-AUG-2000; 2000US-0226779.
 PR 22-AUG-2000; 2000US-0226661.
 PR 22-AUG-2000; 2000US-0226668.
 PR 22-AUG-2000; 2000US-0227109.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228934.
 PR 01-SEP-2000; 2000US-0229587.
 PR 01-SEP-2000; 2000US-0229543.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231142.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231113.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0231480.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0231964.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0233066.
 PR 21-SEP-2000; 2000US-0233067.
 PR 25-SEP-2000; 2000US-0234957.
 PR 25-SEP-2000; 2000US-0234958.
 PR 26-SEP-2000; 2000US-0235084.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0239355.
 PR 13-OCT-2000; 2000US-0239357.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241800.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246533.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249254.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249219.
 PR 17-NOV-2000; 2000US-0249245.
 PR 01-DEC-2000; 2000US-0251060.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0251060.
 PR 08-DEC-2000; 2000US-0250391.
 PR 08-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0249189.
 PR 05-DEC-2000; 2000US-0249179.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Rosen CA, Barash SC, Ruben SM;
 XX DR WPI; 2001-465570/50.
 XX N-PSDB; AAI02073.
 XX PT Isolated nucleic acid molecule encoding a reproductive system antigen is used in preventing, treating or ameliorating a medical condition -
 XX PS Claim 11; SBQ ID NO 4761; 1297pp + Sequence Listing; English.
 XX CC The present invention provides the protein and coding sequences of a number of human reproductive system related antigens. These can be used in the prevention and treatment of reproductive system disorders, including cancer. The present sequence is a protein of the invention.
 XX SQ Sequence 136 AA;
 Best Local Similarity 28.4%; Pred. No. 1.1;
 Score 79; DB 22; Length 136;

Matches 33; Conservative 5; Mismatches 36; Indels 42; Gaps 5;
Qy 7 PMPCAHPLGLFLGLHPALSLPLVTVAGYMSAT-----PRHGLQCPRAAPPAYTGETG 61
Db 33 PAPC-----LGGVHXDFQ--VGPPGIFTSAADPSPWPGCHNLHQTGPRAPPVS--- 78
Qy 62 DSGARETVSDKRSQGH-----TWCTULPRHWTWIG 94
Db 79 ESWPYPLTSQVSPSYSHMDVYMRHTLMPCTTATITITLILALPWIPMG 134

Search completed: June 17, 2003, 10:34:57
Job time : 71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 03:24:38 ; Search time 1031 Seconds
(without alignments)
8480.466 Million cell updates/sec

Title: US-09-955-807-1_COPY_59_373
Perfect score: 315
Sequence: 1 atgtcggttattcttgagcc.....atgtgtttcaqcgagccac 315

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Genbank:*

Result No.	Score	Query Match	Length	DB .ID	Description
1	313.4	99.5	158219	9 AC005534	AC005534 Homo sapi
2	38.8	12.3	136124	9 HS394P21	AL021528 Homo sapi
3	38.8	12.3	155112	2 AL357521	AL357521 Homo sapi
4	38.8	12.3	158297	2 AL513545	AL513545 Homo sapi
5	38.8	12.3	209676	2 AL831755	AL831755 Homo sapi
6	37	11.7	98256	2 AC126501	AC126501 Rattus no
7	36.8	11.7	15666	2 AC105744	AC105744 Oryza sat
8	36.6	11.6	116704	10 AL591921	AL591921 Mouse DNA
9	36.6	11.6	192153	2 AC095187	AC095187 Rattus no
10	36.6	11.6	209882	2 AC10554	AC10554 Rattus no
11	36.2	11.5	3688	10 RN07628	AJ007628 Rattus no
12	36.2	11.5	3736	6 AR179196	AR179196 Sequence
13	36.2	11.5	3736	10 AB026599	AB026599 Rattus no
14	36.2	11.5	10732	6 E3986	E32986 Gene encodi
15	36.2	11.5	110000	2 AC021632_1	Continuation (2 of
16	35.2	11.5	142802	2 AC127725	AC127725 Rattus no
17	36.2	11.5	145616	9 HS108K11	HS108K11 Human DNA
18	36.2	11.5	152992	2 AC023804	AC023804 Mus muscu
19	36.2	11.5	16881	10 AL591459	AL591459 Mouse DNA
20	36.2	11.5	263776	2 AC087335	AC087335 Mus muscu
21	35.4	11.2	43012	9 AC079891	AC079891 Homo sapi
22	35.4	11.2	11253	9 HS273F20	AC127725 Rattus no
23	35.4	11.2	185126	2 AC16269	AC16269 Human DNA
24	35.4	11.2	29267	2 AC023833	AC023833 Homo sapi
25	35.2	11.2	300233	2 AC123842	AC123842 Mus muscu
26	35	11.1	164290	2 AC096454	AC096454 Rattus no
27	35	11.1	191015	2 AC0977557	AC0977557 Rattus no
28	34.8	11.2	99477	9 AL590438	AL590438 Human DNA
29	34.8	11.2	18526	2 AC154743	AC154743 Human DNA
30	34.8	11.0	196329	2 AC129673	AC129673 Rattus no
31	34.8	11.0	154441	2 AC019182	AC019182 Homo sapi
32	34.6	11.0	16443	2 AC011766	AC011766 Homo sapi
33	34.6	11.0	162133	2 AC068569	AC068569 Homo sapi
34	34.6	11.0	174302	2 AC022692	AC022692 Homo sapi
35	34.4	10.9	140572	9 AL354743	AL354743 Human DNA
36	34.4	10.9	128332	9 AC068657	AC068657 Homo sapi
37	34.4	10.9	202559	10 AL645849	AL645849 Mouse DNA
38	34.4	10.9	218904	2 AL844840	AL844840 Mus muscu
39	34.4	10.9	240236	2 AC124765	AC124765 Mus muscu
40	34.2	10.9	51953	2 AC083379	AC083379 Homo sapi
41	34.2	10.9	161355	9 HSJD60019	AL080314 Human DNA
42	34.2	10.9	184365	2 AC079639	AC079639 Mus muscu
43	34	10.8	109832	2 AC120978	AC120978 Rattus no
44	34	10.8	150591	2 AC131143	AC131143 Rattus no
45	34	10.8	157197	2 AC099042	AC099042 Oryza sat

ALIGNMENTS

RESULT 1
AC005534

LOCUS AC005534 158219 bp DNA linear PRI 21-DEC-1999
DEFINITION Homo sapiens PAC clone RP5-982E9 from 7q35-q36, complete sequence.
ACCESSION AC005534
VERSION AC005534.2 GI:4753272
KEYWORDS HTG
HTG
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 158219)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE 99063792
PUBMED 9947074

REFERENCE 2 (bases 1 to 158219)

AUTHORS Drone,K., Wohldmann,P. and Eldred,J.

TITLE The sequence of Homo sapiens PAC clone RP5-982E9

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 158219)

AUTHORS Waterston,R.H.

TITLE direct Submission

JOURNAL Submitted (25-AUG-1998) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 158219)

AUTHORS Waterston,R.H.

TITLE direct Submission

JOURNAL Submitted (05-MAY-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 158219)

AUTHORS Waterston,R.

TITLE direct Submission

JOURNAL Submitted (10-JUL-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE 6 (bases 1 to 158219)

AUTHORS Waterston,R.

TITLE direct Submission

JOURNAL Submitted (21-DEC-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

COMMENT On MAY 5, 1999 this sequence version replaced g1:3007500.

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: saplins@watson.wustl.edu

Center project name: H_DJ0882E09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.ncbi.nlm.nih.gov/DIR/GRB/CHR7/>, send mailto:egreen@nigri.nih.gov, or see <http://genome.wustl.edu/gsc>.

SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-5, prepared by Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>) using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong. VECTOR: pCPAC2.

NEIGHBORING SEQUENCE INFORMATION:

actual end is at 158219 of RP5-982E9.

A transposon was identified in the cloning vector for RP5-982E9.
FEATURES source
Location/Qualifiers
1. .158219
/organism="Homo sapiens"
/db_xref="taxon:9006"
/chromosome="7"
/map="7q35-q36"
/clone="RP5-982E9"
/clone_id="RPCI-5"
repeat_region
1. .61
/rpt_family="Alu"
1033. .1134
/rpt_family="L1"
1419. .1555
/rpt_family="L1"
1559. .1634
/rpt_family="L1"
1647. .1950
/rpt_family="Alu"
2463. .2573
/rpt_family="L2"
3108. .3253
repeat_region
3338. .3883
/rpt_family="L1"
3879. .5566
/rpt_family="L1"
5567. .5869
/rpt_family="Alu"
5870. .5959
/rpt_family="L1"
5947. .6693
/rpt_family="L1"
6654. .7001
repeat_region
7002. .8787
/rpt_family="L1"
8787. .11323
repeat_region
11324. .11620
/rpt_family="Alu"
11621. .13250
/rpt_family="L1"
13492. .13986
/rpt_family="L1"
14003. .14189
/rpt_family="L1"
14301. .15249
/rpt_family="L1"
15379. .1564
/rpt_family="L1"
15690. .15721
repeat_region
15693. .15721
/rpt_family="L1"
1537. .16069
repeat_region
1537. .16069
/rpt_family="MalR"
16070. .16349
repeat_region
16353. .17739
/rpt_family="L1"
17740. .18883
repeat_region
18884. .19913
/rpt_family="MalR"
20463. .21155
/rpt_family="L1"

AL337521 155112 bp DNA linear HTGS 10-JUL-2001
 LOCUS DEFINITION Homo sapiens chromosome 1 clone RP5-9256022, *** SEQUENCING IN
 PROGRESS **, 20 unordered Pieces.
 ACCESSION AL337521
 VERSION 15521.11 GI:11321997
 KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
 SOURCE Organism
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 155112)
 AUTHORS McLay, K.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK E-mail enquiries: humquery@sanger.ac.uk Clone
 requests: clonerequest@sanger.ac.uk
 On Nov 23, 2000 this sequence version replaced 91:10039960.
 COMMENT ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: d1926022
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: M13; M7815; 7% of reads
 Sequencing vector: plasmid; L08752; 92% of reads
 Chemistry: Dye-terminator Em-amersham; 12% of reads Chemistry:
 Dye-terminator Big Dye; 68% of reads
 Chemistry: dye-terminator ABI; 19% of reads
 Consensus quality: 14025 bases at least Q10
 Consensus quality: 148804 bases at least Q30
 Consensus quality: 151286 bases at least Q20
 Insert size: 15312; sum-of-contigs
 Insert size: 121507; 87.0% error; agarose-^{3P}
 Quality coverage: 5.16x in Q20 bases; sum-of-contigs Quality
 coverage: 7.34x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 20 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 8488: contig of 8488 bp in length
 * 8489 8588: gap of 100 bp
 * 14083: contig of 5495 bp in length
 * 14084 14183: gap of 100 bp
 * 14184 16498: contig of 2315 bp in length
 * 16499 16598: gap of 100 bp
 * 16599 23798: contig of 7200 bp in length
 * 23799 23898: gap of 100 bp
 * 36663 36662: contig of 12764 bp in length
 * 36663 36762: gap of 100 bp
 * 36763 45878: contig of 9116 bp in length
 * 45873 45978: gap of 100 bp
 * 45979 50067: contig of 4089 bp in length
 * 50068 50167: gap of 100 bp
 * 56588: contig of 6421 bp in length
 * 56589 56688: gap of 100 bp
 * 56689 63582: contig of 6894 bp in length
 * 63583 63682: gap of 100 bp
 * 63683 68434: contig of 4752 bp in length
 * 68435 68534: gap of 100 bp
 * 68535 75603: contig of 7069 bp in length
 * 75604 75703: gap of 100 bp
 * 75704 78066: contig of 2358 bp in length
 * 78062 78161: gap of 100 bp
 * 78162 88594: contig of 10793 bp in length

 FEATURES source
 misc_feature /organism="Homo sapiens"
 /db_xref="Taxon:9606"
 /chromosome="1"
 /clone="RP5-9256022"
 /clone_id="RPCI-5"
 /note="assembly_fragment:00871
 /clone_end_sp6
 /vector_end_sp6
 /vector_side:left"
 misc_feature 1..8488
 /note="assembly_fragment:01703
 fragment_chain:1"
 misc_feature 14184..16498
 /note="assembly_fragment:02705
 fragment_chain:1"
 misc_feature 16599..23798
 /note="assembly_fragment:00068
 fragment_chain:1"
 misc_feature 23899..36662
 /note="assembly_fragment:02777
 fragment_chain:1"
 misc_feature 36763..45878
 /note="assembly_fragment:00748
 fragment_chain:2"
 misc_feature 45979..50067
 /note="assembly_fragment:00024
 fragment_chain:2"
 misc_feature 50168..56588
 /note="assembly_fragment:02548
 fragment_chain:2"
 misc_feature 56689..63582
 /note="assembly_fragment:01501
 fragment_chain:3"
 misc_feature 63683..68434
 /note="assembly_fragment:01410
 fragment_chain:3"
 misc_feature 68535..75603
 /note="assembly_fragment:00219"
 misc_feature 75704..78061
 /note="assembly_fragment:00724"
 misc_feature 78162..88594
 /note="assembly_fragment:01966"
 misc_feature 89055..92586
 /note="assembly_fragment:01996"
 misc_feature 92687..96459
 /note="assembly_fragment:02016"
 misc_feature 96560..101475
 /note="assembly_fragment:02131"
 misc_feature 101576..11520
 /note="assembly_fragment:02503"
 misc_feature 11521..11620
 /note="assembly_fragment:02059
 fragment_chain:4
 clone_end:17
 vector_side:right"
 misc_feature 146633..151919
 /note="assembly_fragment:00729

```

*misc_feature
  fragment_chain:4"
  /note="Assembly fragment:01146
  BASE COUNT 38677 a 38785 c 38240 g 37496 t 1914 others
  ORIGIN
    Query Match 12.3%; Score 38.8; DB 2; Length 155112;
    Best Local Similarity 57.4%; Pred. No. 0.34;
    Matches 70; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
    Oy 77 CTTGGCCCGCTGTAGTACCTGGGCTGGAGTGATGAGGCCCACTCCAAAGCATGGCTCG 136
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Db 88691 CGTGGTCCCGTTATGTGATGCTGGTTCTGCAATTCTATAGAACAGGGTCT 88750
    Oy 137 AACATGTGTCCTCCGCCCTCAGCAGAGTCACTGGGACTGGGGCA 195
    Db 88751 CAAATGTGACACCACCCACACAGCAGTGCGCACATCAGGAACTTGCGGGA 88810
    Qy 197 AG 198
    Db 88811 TG 88812

RESULT 4
AL513545 A1513545 158297 bp DNA linear HTG 10-JUL-2001
DEFINITION Homo sapiens chromosome 1 clone RP11-514C23, *** SQUENCING IN
PROGRESS ***, 18 unordered pieces.
ACCESSION AL513545
VERSION A1513545.2 GI:13160538
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE
ORGANISM Homo sapiens
REFERENCE
  Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
  I (bases 1 to 158297)
AUTHORS
  McLAY, K
TITLE
  Direct Submission
JOURNAL
  Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
  requests: clonerequest@sanger.ac.uk Clone
COMMENT
  On Feb 28, 2001 this sequence version replaced 91:12750921.
  Center: Sanger Centre
  Center code: SC
  Contact: humquery@sanger.ac.uk
  ..... Project Information
  Center project name: ba514c23
  ..... Summary Statistics
  Assembly program: XGP4; version 4.5
  Sequencing vector: plasmid; L0752; 100% of reads
  Chemistry: Dye-Terminator Big Dye; 100% of reads
  Consensus quality: 153889 bases at least Q40
  Consensus quality: 15599 bases at least Q30
  Insert size: 15597; sum-of-contigs
  Insert size: 16270; 3.0% error
  Quality coverage: 6.58x in 020 bases; sum-of-contigs
  Coverage: 6.38x in 020 bases; agarose-fp
  ..... This is a 'working draft' sequence. It currently
  * consists of 18 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  * 1. 4316: contig of 4316 bp in length
  * 4317 4416: gap of 100 bp
  * 4417 13709: contig of 9293 bp in length

```

```

FEATURES
  source
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="1"
    /clone="RP11-514C23"
    /clone.lib="RPCI-11.2"
    /note="Assembly fragment:02586
    fragment_chain:1
    clone.end:SP6
    vector_side:left"
    4417..13709
    /note="Assembly fragment:01913
    fragment_chain:1"
    13810..16011
    /note="Assembly fragment:01793
    fragment_chain:1"
    16112..19225
    /note="Assembly fragment:02161
    fragment_chain:1"
    19326..24319
    /note="Assembly fragment:00039
    fragment_chain:1"
    24420..32439
    /note="Assembly fragment:01183
    fragment_chain:1"
    32540..35376
    /note="Assembly fragment:01645
    fragment_chain:1"
    35477..48033
    /note="Assembly fragment:01800
    fragment_chain:2"
    misc_feature
      48134..65465
      /note="Assembly fragment:02272
      fragment_chain:2"
      65567..70347
      /note="Assembly fragment:02380
      fragment_chain:2"
      misc_feature
        70448..87906
        /note="Assembly fragment:00477

```

fragment_chain:2"
88007. .96207
/note="assembly fragment:00137
misc_feature
/note="assembly fragment:01506
98883. .10526
/note="assembly fragment:00660
fragment_chain:3"
96308. .9882
/note="assembly fragment:01627. .130171
/note="assembly fragment:00061
fragment_chain:3"
130272. .140166
/note="assembly fragment:01928
fragment_chain:3"
140267. .132701
/note="assembly fragment:01044.
fragment_chain:3
clone_end:T/
vector_side:right"
BASE COUNT 39009 a 38793 c 40866 g 37912 t 1717 others
ORIGIN
Query Match 12.3%; Score 38.8; DB 2; Length 158297;
Best Local Similarity 57.4%; Pred. No. 0.34; Gaps 0;
Matches 70; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 77 CTTGGCCCTGTAGTACTGAGCTGGAGTGTAGGGCCACTCCAGATGCCCTGG 136
Db 46294 CGGGTCCTGTATTCTGAGCTGCTGTTTCTGCATTCTATAGAACAGTGCTCT 46353
QY 137 AACATGTCCTCGCCCTCCACASAGTGCAGGTTCACTGGGAACTGGGGAA 196
Db 46354 CAATTGGAACCAACCCACACAGCAGTGGCAGCACTGGGAACCTTGCGAA 46413
QY 197 AG 198
Db 46414 TG 46415

RESULT 5
AL831755
LOCUS AL831755 209676 bp DNA linear HTG 09-AUG-2002
DEFINITION Homo sapiens chromosome 1 clone RP13-279N23, *** SEQUENCING IN PROGRESS
PROGRESS ***, in ordered pieces.
ACCESSION AL831755
VERSION GI:2220471
KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 209676)
AUTHORS Harrison,E.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK. E-mail enquiries: humquer@sanger.ac.uk Clone requests: cloneresquests@sanger.ac.uk On Aug 11, 2002 this sequence version replaced g1:2198292. Genome Center Center: Wellcome Trust Sanger Institute Center code: SC Web site: http://www.sanger.ac.uk Contact: humquer@sanger.ac.uk Project Information Center Project name: bb279N23 Summary Statistics Assembly program: XMAP4; version 4.5 Sequencing vector: Plasmid; I08752; 100a of reads

RESULT 6
AC126501
LOCUS AC126501 98256 bp DNA linear HTG 24-JUL-2002
DEFINITION Rattus norvegicus clone CH230-206K3, *** SEQUENCING IN PROGRESS
PROGRESS ***, 30 unordered pieces.
ACCESSION AC126501
VERSION AC126501.2 GI:21703477
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 98256)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amaraluna,H.C., Are,J.R., Ayalew,M., Banks,T., Barbala,J., Bentton,J., Blimage,K., Blankenburg,K., Bonin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buahy,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Ehrhart,C., Edgar,D., Edwards,C.C., Elhaij,C., Escott,M., Farris,T., Ferragut,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Chemistry: Dye-terminator; 97% of reads
Chemistry: dye-terminator Big Dye; 2% of reads
Consensus quality: 209220 bases at least 040
Consensus quality: 209374 bases at least 030
Consensus quality: 209478 bases at least 020
Insert size: 20959; 3.2x error; agarose-fp
Quality coverage: 12.24x in Q20 bases; sum-of-contigs Quality
coverage: 11.72x in Q20 bases; agarose-fp
.....
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
the accession number will be preserved.
Location/Qualifiers
1. .209676
/note="assembly fragment:03689
/organism="Homo sapiens"
/ab_xref="taxon:9606"
/chromosome="1"
/clone="RPCI-27N23"
/clone.lib="RPCI-13.2"
1. .209676
/note="assembly fragment:03689
/note="assembly fragment:03689
clone_end:sp6
vector_side:right"
BASE COUNT 52994 a 53379 c 52820 g 50483 t
ORIGIN
Query Match 12.3%; Score 38.8; DB 2; Length 209676;
Best Local Similarity 57.4%; Pred. No. 0.34; Gaps 0;
Matches 70; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 77 CTTGGCCCTGTAGTACTGAGCTGGAGTGTAGGGCCACTCCAGATGCCCTGG 136
Db 28788 CGGGTCCTGTATTCTGCAGTACCTGGGAACTGGGGAA 28847
QY 137 AACATGTCCTCGCCCTCCACAGCAGTGCAGGTTCACTGGGAA 196
Db 28848 CAATTGGAACCAACCCACACAGCAGTGGCAGCATCTGGGAACCTTGCTGGAA 28907
QY 197 AG 198
Db 28908 TG 28909

ZHAO, S.; DUNN, D.; VON NIEDERHAUERN, A.; WEISS, R.; SMITH, D.R.,
 HOLT, R.A.; SMITH, H.O.; WEINSTOCK, G. and GIBBS, R.A.; SMITH, D.R.,
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 REFERENCE
 AUTHORS
 COMMENT
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 REFERENCE
 AUTHORS
 COMMENT
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

Center project name: GGF

Center clone name: CH20-120K20

Summary Statistics

Sequencing vector: plasmid;
 Chemistry: dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.90329
 Consensus quality: 167910 bases at least Q40
 Consensus quality: 178487 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length.
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank/draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 60 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1. 1129: contig of 1129 bp in length
 * 1130 1229: gap of unknown length
 * 1230 3065: contig of 1836 bp in length
 * 3066 3165: gap of unknown length
 * 3166 4192: contig of 1027 bp in length
 * 4193 4292: gap of unknown length
 * 4293 5851: contig of 1559 bp in length
 * 5852 5951: gap of unknown length
 * 5952 7127: contig of 1176 bp in length
 * 7128 7228: gap of unknown length
 * 7228 8256: contig of 1029 bp in length
 * 8257 8356: gap of unknown length
 * 8357 9467: contig of 1111 bp in length
 * 9468 9567: gap of unknown length
 * 9568 10838: contig of 127 bp in length
 * 10839 10938: gap of unknown length
 * 10939 12068: contig of 1130 bp in length
 * 12069 12168: gap of unknown length
 * 12169 13196: contig of 1028 bp in length
 * 13197 13296: gap of unknown length
 * 13297 14315: contig of 1011 bp in length
 * 14316 14416: gap of unknown length
 * 14416 15738: contig of 1323 bp in length
 * 15739 15838: gap of unknown length
 * 15839 17887: contig of 2049 bp in length
 * 17888 17987: gap of unknown length
 * 17989 19728: contig of 1741 bp in length
 * 19729 19828: gap of unknown length
 * 19829 21303: contig of 1475 bp in length
 * 21304: gap of unknown length

21404 22478: contig of 1275 bp in length
 22679 22778: gap of unknown length
 22779 23055: contig of 1177 bp in length
 23055 24055: gap of unknown length
 24056 26396: contig of 2241 bp in length
 26396 26397: gap of unknown length
 26397 27535: contig of 1139 bp in length
 27536 27635: gap of unknown length
 27636 29882: contig of 1947 bp in length
 29882 30507: gap of unknown length
 30507 31852: contig of 2170 bp in length
 31852 31952: gap of unknown length
 31952 33524: contig of 1572 bp in length
 33525 33624: gap of unknown length
 33625 3607: contig of 2783 bp in length
 3607 40998: gap of unknown length
 40998 43389: contig of 2291 bp in length
 43389 43390: gap of unknown length
 43390 45173: contig of 1784 bp in length
 45173 45273: gap of unknown length
 45274 47552: contig of 2046 bp in length
 47553 49552: gap of unknown length
 49552 49557: contig of 1805 bp in length
 49557 52516: contig of 2959 bp in length
 52516 52616: gap of unknown length
 52616 55579: contig of 2963 bp in length
 55580 55579: gap of unknown length
 55580 55679: gap of unknown length
 55680 57710: contig of 2031 bp in length
 57711 57810: gap of unknown length
 57811 60388: contig of 2578 bp in length
 60388 60489: gap of unknown length
 60489 63297: contig of 2809 bp in length
 63297 63397: gap of unknown length
 63398 67352: contig of 3955 bp in length
 67352 67452: gap of unknown length
 67452 71221: contig of 3769 bp in length
 71221 71321: gap of unknown length
 71322 74908: contig of 3587 bp in length
 74909 75009: gap of unknown length
 75009 78851: contig of 3843 bp in length
 * 78852 78951: gap of unknown length
 * 78952 81039: contig of 2088 bp in length
 * 81040 81139: gap of unknown length
 * 81140 84921: contig of 3782 bp in length
 * 84922 85021: gap of unknown length
 * 85022 90311: contig of 5290 bp in length
 * 90311 90411: gap of unknown length
 * 90412 94525: contig of 4114 bp in length
 * 94526 94625: gap of unknown length
 * 94626 98229: contig of 3604 bp in length
 * 98230 98329: gap of unknown length
 * 98330 102008: contig of 3679 bp in length
 * 102009 102108: gap of unknown length
 * 102109 106715: contig of 4607 bp in length
 * 106716 106815: gap of unknown length
 * 106815 106900: contig of 4085 bp in length
 * 106900 111000: gap of unknown length
 * 111001 115834: contig of 4834 bp in length
 * 115834 115934: gap of unknown length
 * 115935 119585: contig of 3651 bp in length
 * 119586 119685: gap of unknown length
 * 119686 124375: contig of 4690 bp in length
 * 124376 128981: contig of 4506 bp in length

Query Match 11.6%; Score 36.6; DB 2; Length 209882;
 Best Local Similarity 52.3%; Pred. No. 1.7;
 Matches 81; Conservative 0; Mismatches 74; Index 0; Gaps 0;

Mon Jun 23 10:42:32 2003

us-09-955-807-1_copy_59_373.rge

Page 15

Db 33562 GGGCTCCCTGGCTCTGCGCTGGGTAGCAGCTGCCAGCTGGACTGG 33514

Search completed: June 21, 2003, 03:52:44
Job time : 1086 secs

THIS PAGE BLANK (USPTO)

Gencore version 5.1.6
copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 17, 2003, 10:31:01 ; Search time 80 seconds

(without alignments) (270.437 Million cell updates/sec)

Title: US-09-955-807-2

Perfect score: 582

Sequence: MIGYSBMPCAHPLGLFLLG.....PHPWLTWVGHJINHYSSASH 105

Scoring table: BL0SUM62

Gapop 10.0 , Gapext: 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: STREMLI_21.*
- 2: sp_archea:*
- 3: sp_bacteria:*
- 4: sp_fungi:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp Rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	89	15.3	678	5	Q960P9 drosophila
2	89	15.3	743	5	Q9VBW6 drosophila
3	79.5	13.7	684	4	Q9H6M8 drosophila
4	79.5	13.7	684	4	Q96BD1 homo sapien
5	79.5	13.7	684	4	Q9FSK3 oryza sativa
6	79	13.6	261	10	Q9AUN2 oryza sativa
7	76.5	13.1	206	10	Q8RUU7 oryza sativa
8	76.5	13.1	160	5	Q9NKN0 leishmania
9	75.5	13.0	549	13	Q9w639 xenopus laevis
10	74	12.7	220	11	Q9CVP2 mus musculus
11	73.5	12.6	132	2	Q9RPQ0 chlamydia trachomatis
12	73.5	12.6	357	4	Q9NZL4 homosapiens
13	73	12.5	135	9	Q9MC96 bacteriophage
14	72.5	12.5	354	12	Q9WRN7 macaca mulatta
15	72.5	12.5	554	10	Q26911 arabidopsis thaliana
16	72	12.4	1089	16	Q9X908 streptomyces

ALIGNMENTS

17	71	12.2	110	2	Q53356 streptomyces
18	71	12.2	494	10	Q86ab8 oryza sativa
19	71	12.2	1217	4	Q9u115 homo sapiens
20	70	12.0	497	16	Q93JB9 streptomyces
21	70	12.0	679	10	Q9FV53 arabidopsis
22	70	12.0	1620	16	Q96285 mycobacterium
23	69.5	11.9	141	4	Q8TEG2
24	69.5	11.9	147	4	Q9HSU7
25	69.5	11.9	434	16	Q9RVL8
26	69.5	11.9	860	16	Q9PEZ1
27	69.5	11.9	883	2	Q5102 streptomyces
28	69	11.9	287	5	Q20754 caenorhabditis
29	69	11.9	328	16	Q9A7U1
30	69	11.9	365	5	Q8T4A4
31	69	11.9	447	5	Q9VLJ5
32	69	11.9	507	13	Q9WTK3
33	69	11.9	588	10	Q9SDC0
34	68.5	11.8	194	5	P91787 drosophila
35	68.5	11.8	202	16	Q55623 synochocystis
36	68.5	11.8	323	11	Q9EOW6
37	68.5	11.8	68.5	2	Q9F8U5 streptomyces
38	68.5	11.8	399	16	Q8YF50 anabaena sp
39	68.5	11.8	1235	5	Q9VFV3
40	68	11.7	586	5	Q9W1Y2
41	68	11.7	1400	16	P96419 mycobacterium
42	67.5	11.6	220	4	Q96SJ1
43	67.5	11.6	4	Q96ID9	
44	67.5	11.6	333	4	Q96NT7
45	67.5	11.6	352	4	Q9BYE7

RESULT 2		RESULT 3	
O9VBW6	PRELIMINARY;	PRT;	743 AA.
ID O9VBW6			
AC O9VBW6;			
DT 01-MAY-2000 (TREMBLrel. 13, Created)			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DE CG11849 protein.			
GN Drosophila melanogaster (Fruit fly).			
OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; OC			
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; OC			
OC Ephydidae; Drosophilidae; Drosophila. OC			
OX NCBI_TAXID=7227; OX			
RN [1]			
RP STRAIN-BERKELEY;			
RX SPROQUENCE FROM N.A.			
MLDLINE=20196006; Pubmed=10731132;			
RA M. D., Celinkin, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., RA Amannides, P.G., Scheerer, S.E., Li, P.W., Hoskins, R.A., Gallo, R.F., RA Sutton, G.A., Lewitus, S.E., Richards, S., Asburner, M., Henderson, S.N., RA Brandon, R.C., Wortsman, J.R., Yandell, M.D., Zhang, Q., Chen, L.X., RA Rogers, Y.-H.C., Blazej, R.G., Champé, M., Pfeiffer, B.D., RA Wan, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Miklos, G.L.G., RA Abel, J.F., Agbayani, A., An, H.-J., Andrews, P.-F., Fanckoch, C., Baldwin, D., RA Belilew, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., RA Borsova, K.Y., Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., RA Borkson, D., Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P., RA Burtsis, K.C., Busam, D.A., Butler, H., Cadieu, E., Centner, A., Chandra, I., RA Cherry, J.M., Cawley, S., Danilke, C., Davenport, L.B., Davies, P., RA De Pablo, B., Deicher, A., Deng, Z., Mays, A.D., Dew, T., Dietz, S.M., RA Dodson, K., Douc, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., RA Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferreira, S., Fleischmann, W., RA Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., RA Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., RA Hostin, D., Houston, K.A., Howland, T.J., Wei, M.-H., Ibegwam, C., RA Jalali, M., Kalush, F., Karpen, G.H., Ke, Z., Kennison, J.A., Ketchum, K.A., RA Kummel, B.E., Kodira, C.D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., RA Lasko, P., Lei, Y., Levitt, J., Li, J., Li, Z., Liang, Y., Lin, X., RA Liu, X., Matteei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., RA Merkulov, G., Mishina, N.V., Morris, J.J., Mosheri, A., RA Mount, S.M., Moy, M., Murphy, B., Muny, D.M., Nelson, D.L., RA Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Paclet, J.M., RA Palazzolo, M., Pittman, G.S., Pan, S., Pollard, J.J., Purk, V., Reese, M.G., RA Reinert, K., Remington, K., Saunders, R.D.C., Schaefer, F., Shen, H., RA Shue, B.C., Siden-Kiamios, I., Simpson, M., Skupski, M.P., Smith, T., RA Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, B., RA Svistsk, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., RA Wang, Z.-Y., Wasserman, D.A., Weinstock, G.M., Weissbach, J., RA Williams, S.M., Woodage, T., Worley, K.C., Wu, D., Yang, S., Yao, Q.A., RA Ye, J., Yeh, R.-F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., RA Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., RA Myers, E.W., Robin, G.M., Venter, J.C.; RT The genome sequence of Drosophila melanogaster.; RT Science 287: 2185-2195 (2000). DR EMBL; AE003751; AA054411.; DR FLYbase; FBgn0039286; CG11849. SQ SEQUENCE: 743 AA; 78863 MW; B32C5DBD499B8F6 CRC64;			
Query Match 15.3%; Score 89; DB 5; Length 743;			
Best Local Similarity 32.4%; Pred. No. 0 44;			
Matches 22; Conservative 15; Mismatches 23; Indels 8; Gaps 3;			
OY 18 LGGLHPL--SPLWVTVAGVMSATPKH--GEQCPAPPAPPVFGFDGAKETVSQDK 73	PRELIMINARY;	PRT;	684 AA.
Db 330 MGGLNPFLPPPTIPMGVVISPPINSSTPQHMSQLAQTPPPIPSAPLTPSSMPGSLD---EK 385			
OY 74 ROGHTWC 81			
Db 386 INAWYNNC 393			
Query Match 13.7%; Score 79.5; DB 4; Length 684;			
Best Local Similarity 29.7%; Pred. No. 4; Matches 35; Conservative 13; Mismatches 53; Indels 17; Gaps 5;			
OY 2 LGYSEPMCAHPIGFL-LGGLHPLS-----LVTVAGVMSATPKHGLEQ-----P 49			

Db	1 MGLSAAAPLWGGPPGILLALAHPLASVPRRDCVLGAGPAGLQMAVFLORGARDYAVFE 60	DR InterPro; IPR005213; HGWP.
QY	50 PAPPAPAVIGFTGDSGAKETIVSODKRSQGHTWTQTLAHPWLWGH---LRNHYSSA 103	DR Pfam; PF03578; HGWP; 2.
Db	61 RAPRG-SIFFRYPRHRKLISINKRYTCKANEFLNRHDWSNSLSDHORLIERHNSRA 117	KW Hypothetical protein.
RESULT 5		SEQUENCE 261 AA; 28496 MW; DEC11EC16B4B05B CRC64;
Q9FSK3	PRELIMINARY; PRT; 826 AA.	Query Match Best Local Similarity 26.5%; Score 79; DB 10; Length 261; Matches 36; Conservative 9; Mismatches 43; Indels 48; Gaps 5;
ID Q9FSK3		DR 12 HPLGLPFLGLHPALSPLPVTVAGVMS---ATPKHGLEQQCPPAPPVTRGFGDS---- 63
AC 09FSK3:		AC 4 HRAGVILGLGAQSCLVPGPVAVGGVGSVLLSMARMGWICHVVERPYAVWVMSDQSTYS 63
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)		Db 64 -----GAKETIVSODKRSQG-----HTWC-----TIAUPHPWLWV- 93
DT 01-MAR-2001 (TREMBLrel. 19, Last annotation update)		QY 64 NSRALGYGQAOCTQEWIONITGQLSPRLHSWLVTPPLGVVIFTAGLSSPLITCQFA 123
DE Hypothetical 92.0 kDa protein.		Db H08605.17.
GN Oryza sativa (Rice).		QY 94 -----GHLRNH 99
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Elhardtioideae; Oryzeae.		DR 124 VMAIGHTAVDGHRLH 139
OC NCBI_TAXID=4530;		RN
RN		SEQUENCE FROM N.A.
RP		Han B., Feng Q., Huang Y.C., Chen Z.H., Zhou B., Li Y., Zhu J.J., Tang Y.S., Zhao Q., Liu Y.L., Mu J., Yu Z., Fan D.L., Chen L., Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Zhu J., Liu X.H., Hu X., Lei H.Y., Zhang Y.J., Wang R., Li C., Lu Y., Chen X.C., Zhang Y., RA RA RA RA RT "Oryza sativa indica (Guanglu14) genomic DNA, chromosome 4, BAC clone: H08605," Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RT RL DR EMBL; AT4A2113; CAC09484.1; -.		DR 14 PLPLHP-----IHPQSLAAVASLADPTEVVAAPACGSSGLPPPPQQA 66
KW HYPOTHETICAL PROTEIN		DE HYPOTHEICAL 22.5 kDa protein.
SQ SEQUENCE 826 AA; 92046 MW; 122951E5B874E6A8 CRC64;		GN OSJNBA0091J06.11 OR oil34F06.4.
[1]		DR 56 VTFGFDGSGAKETIVSODKRSQGHT-----WCTL-----ALPHPHWLW 92
RESULT 6		DR 67 SP-TMDSTAASPSDPKQARWPTMAGDGRWVMDIWHNLUGDAVGNLWSSW 118
O9AUN2	PRELIMINARY; PRT; 261 AA.	RN
ID O9AUN2		SEQUENCE FROM N.A.
AC O9AUN2;		DR STRAIN=CV_NIPONBARE;
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)		RC 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)		RC 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)		DR HYPOTHETICAL 28.5 kDa protein.
DR H08605.17.		GN Oryza sativa (Japanica cultivar-group).
GN OSJNBA0058E19.19.		OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Elhardtioideae; Oryzeae; Oryza.
Db		OC NCBI_TAXID=39947;
RN		DR 11
RP		SEQUENCE FROM N.A.
RC		DR STRAIN=CV_NIPONBARE;
RC		RC 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
RA		RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
RA		RA Kult K., Nasimento L., Spiegel L., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
RA		RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA		RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA		RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.; strain, clone
RT		RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone OSJNBA0091J06, from chromosome 10, complete sequence.";
RT		RT Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RL		RL Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN		[2]
RP		DR 01341F06, from chromosome 10, complete sequence.";
RC		DR Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RA		DR EMBL; AC11338; AAM08655.1; -.
RA		DR EMBL; AC116926; AAM08884.1; -.
RA		DR EMBL; AC116926; AAM08884.1; -.
RN		KW HYPOTHETICAL PROTEIN.
RP		SEQUENCE FROM N.A.
RC		DR 12 HPLGLPFLGLHPALSPLPVTVAGVMS---ATPKHGLEQCPAPPVTRGFGDS---- 54
RA		RC 11 AHPLGLPFLGLHPALSPLPVTVAG---VMSATPKHGLEQ-----PPAPP 54
RA		RA 14 AHAVHILPGIHPKRSQGHTAFTVGDXLYMEASPERLPWAHCFCFEALIHCIPAPP 72
RA		RA 55 AVIGFTGDSGAKETIVSODKRSQH 73
RA		RA 73 SRTG---TGAPSQRQARRIAAH 92
RT		RT "Genomic Sequence For Oryza sativa, Nipponbare strain, Chromosome X,"
RT		RT Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RL		DR EMBL; AC083945; AAK13144.1; -.
DR		

RESULT 8		Best Local Similarity 26.7%; Pred. No. 8.4; Matches 27; Conservative 8; Mismatches 37; Indels 29; Gaps 4;					
Q9RNKO	PRELIMINARY;	PRT;	1660 AA.	AC	O9NKK0;	01-OCT-2000 (TREMBrel. 15, Created)	DT DT
Q9RNKO				AC	O9NKK0;	01-OCT-2000 (TREMBrel. 15, Last sequence update)	DT DT
				AC	O9NKK0;	01-DEC-2001 (TREMBrel. 19, Last annotation update)	DE DE
				GN	L5204_7.		GN
OC	Leishmania major.			OC			
NCBI_TaxID=5664;	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.			RN	[1]		
RP	SEQUENCE FROM N.A.			RN			
RC	STRAIN=FREDLIN;			RC			
RA	Myler P.J., Sisk E., Hixson G., Kiser P., Rickel E., Hassebroek M.,			RA			
RL	Cawthra J., Marcolini F., Sunkin S., Stuart K.D.; Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.			RL			
DR	EMBL; AAC05941; AAP34394_1.			DR			
DR	InterPro; IPR002965; P_RICH_extensions.			DR			
DR	InterPro; IPR000217; Tubulin.			DR			
DR	PRINTS; PRO1217; PRICHXTENS.			DR			
DR	PROSITE; PS00227; TUBULIN_UNKOWN_1.			DR			
SQ	SEQUENCE 1660 AA; 170277 MW; 3BCD35A874346116 CRC64;			SQ			
Query Match	Best Local Similarity 34.1%; Score 76.5; DB 5; Length 1660; Matches 31; Conservative 7; Mismatches 36; Indels 17; Gaps 5;			Query Match	13.1%; Score 76.5; DB 5; Length 1660; Matches 31; Conservative 7; Mismatches 36; Indels 17; Gaps 5;		
Db	6 EPMPCAHDLIGFLGLHPALSLPLVWVAGMSATRKHGQCP-----PAPPAT 57			Db	219 PTSQP-----ASLLAATHNDGLSLIAPVPPGQHTONGSAQPTYHNNTTWGSRTA 273		
Db	360 EPBP-ASPVTMPL--LAATPLASPAINTSAAPPRSLCALPLSSASRPPSPPPAV- 414			Db	67 ETVSQDKRSQGHTWCTALPH-----LRNH 99		
Qy	58 GFTGDSGAKTVSQDKRSQGHTWCTALPH 88			Qy			
Db	415 --PSPSSAAPTTA--PASPTRPRCOPARPPP 441			Db	274 YTPNMSSHHPNLQHHPMPHMPGHYWPMPHNELAQPPISNH 314		
RESULT 9							
Q9W639	PRELIMINARY;	PRT;	549 AA.	Q9W639	SEQUENCE FROM N.A.		
AC	09W639;			AC	STRAIN=C57BL/6J; TISSUE=TESTIS;		
DT	01-NOV-1999 (TREMBrel. 12, Created)			DT	MEDLINE=21085660; PubMed=1121851;		
DT	01-NOV-1999 (TREMBrel. 12, Last sequence update)			DT	Kawai J., Shihabawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Konio S., Yamamoto I., Saito T., Okazaki Y., Gotobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nakaido T., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Camincci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Sevat T., Shibata Y., Storch K.-F., Suzuki H., Tovo-H., Wang K.-H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.,		
DE	01-JUN-2002 (TREMBrel. 21, Last annotation update)			DE	"Functional annotation of a full-length mouse cDNA collection.", Nature 409:695-699(2000).		
GN	XSMAD4A.			GN	Nature 409:695-699(2000).		
OS	Xenopus laevis (African clawed frog).			OS	EMBL; AK007165; BAB24879_1; -		
OC	Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;			OC	MGI; MGI:1920819; 1700110N18RIK.		
OC	Xenopoda; Xenopus.			OC	FT NON_TER 1		
RN	[1]			SQ	SEQUENCE 220 AA; 24067 MW; 9D8780481470406A CRC64;		
RP	SEQUENCE FROM N.A.			Query Match	12.7%; Score 74; DB 11; Length 220; Best Local Similarity 27.2%; Pred. No. 4.6; Matches 28; Conservative 28; Mismatches 47; Indels 16; Gaps 4;		
RX	MEDLINE=9922549; Published=10207044;			Query Match	12.7%; Score 74; DB 11; Length 220; Best Local Similarity 27.2%; Pred. No. 4.6; Matches 28; Conservative 28; Mismatches 47; Indels 16; Gaps 4;		
RT	"Identification of two Smad4 proteins in Xenopus: thier common and distinct properties.", J. Biol. Chem. 274:12153-12170(1999).			RT	7 PMPCAHDLIGFLGLHP-AISLPLVWVAGMSATRKHGQCPAPPVAVGFTGSG 64		
RL	EMBL; AB022721; BAA77514_1; -			Db	56 PABC-----LGGVHRDFQVIAHGFTFTADPNWRPGHLHQHGTGPAPPPTAS---ESW 104		
DR	HSSP; Q13495; ID: I04153; -			Db	65 AKETVSQDKRSQGHTWCTALPH-----LRNH 99		
DR	TRANSFAC; I04153; -			Db	105 HYPLASQVSPVSHMDYCRHHPHAHVHRHHHHHPTAGS 147		
DR	InterPro; IPR001132; Dwarfin.			DR			
DR	InterPro; IPR003619; Dwarfin_A.			DR			
DR	InterPro; IPR004863; MHL.			DR			
DR	PRINTS; PRO1217; PRICHXTENS.			DR			
DR	PROSITE; PS003165; MHL; 1.			DR			
DR	SMART; SM00523; DWA; 1.			DR			
DR	SMART; SM00524; DWB; 1.			DR			
SQ	SEQUENCE 549 AA; 59855 MW; 72431FD6AC674150 CRC64;			SQ			
Query Match	13.0%; Score 75.5; DB 13; Length 549;			Query Match	13.0%; Score 75.5; DB 13; Length 549;		
RESULT 10							
Q9CVP2	PRELIMINARY;	PRT;	220 AA.	Q9CVP2	23 PALSLPLVWVAGMSATRKHGQCPAPP-----PA-----VTCFTGSGAK 66		
ID	09CVP2;			ID	Db 219 PTSQP-----ASLLAATHNDGLSLIAPVPPGQHTONGSAQPTYHNNTTWGSRTA 273		
AC	09CVP2;			AC	Qy 67 ETVSQDKRSQGHTWCTALPH-----LRNH 99		
DT	01-JUN-2001 (TREMBrel. 17, Last sequence update)			DT	Db 274 YTPNMSSHHPNLQHHPMPHMPGHYWPMPHNELAQPPISNH 314		
DE	1700110N18RIK protein (Fragment).			DE			
GN	1700110N18RIK			GN			
OS	Mus musculus (Mouse).			OS			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OX NCBI_TaxID=10090;			OC			
RN	[1]			RN			
RP	SEQUENCE FROM N.A.			RP			
RC	STRAIN=C57BL/6J;			RC			
RA	MEIDLINE=21085660; PubMed=1121851;			RA			
RA	Kawai J., Shihabawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Konio S., Yamamoto I., Saito T., Okazaki Y., Gotobori T., Bono H., Kasukawa T., Saito R., Radota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nakaido T., Resole G., Quackenbush J., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Camincci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Sevat T., Shibata Y., Storch K.-F., Suzuki H., Tovo-H., Wang K.-H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.,			RA			
RA	"Functional annotation of a full-length mouse cDNA collection.", Nature 409:695-699(2001).			RA			
RA	Nature 409:695-699(2001).			RA			
RA	MGI; MGI:1920819; 1700110N18RIK.			RA			
RT	FT NON_TER 1			RT			

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE Inclusion membrane protein E.
GN INCE.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN NCBI_TAXID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN="LGV-43";
RX MEDLINE=9377128; PubMed=10447885;
RA Scidmore-Carlson M.A., Shaw E.I., Dooley C.A., Fischer E.R.,
RA Hackstadt T.;
RT "Identification and characterization of a Chlamydia trachomatis early operon encoding four novel inclusion membrane proteins.";
RL Mol Microbiol. 33: 753-765 (1999).
DR EMBL: AF15134; AND43915.1; -.
DR InterPro: IPR000847; HTH_LYR.
DR PROSITE: PS00044; HTH_LYSR_FAMILY; UNKNOWN_1.
SQ SEQUENCE 132 AA; 13594 MW; IA4A9FF78B1E6C1E CRC64;

Query Match 12.6%; Score 73.5; DB 2; Length 132;
Best Local Similarity 34.4%; Pred. No. 3.1;
Matches 22; Conservative 7; Mismatches 26; Indels 9; Gaps 2;
DE Basic helix-loop-helix protein class B 1 (Fragment).
GN BHBL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TAXID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20205664; PubMed=10737801;
RA Wang J., Janni-Sait S.N., Escalon E.A., Carroll A.J., de Jong P.J.,
RA Kirsch I.R., Aplan P.D.;
RT "The t(14;21)(q11.2;q22) chromosomal translocation associated with T-cell acute lymphoblastic leukemia activates the BHBL1 gene.;"
RN Proc. Natl. Acad. Sci. U.S.A. 97:3497-3502(2000).
DR EMBL: AF21520; AFM61215.1; -.
DR InterPro: IPR00102; HHL_basic.
DR InterPro: IPR000408; Reg_chr_condens.
DR Pfam: PF00010; HHL; 1.
DR SMART: SM00355; HHL; 1.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE: PS00626; RCC1_2; UNKNOWN_1.
FT NON_TER 1
FT SEQUENCE 357 AA; 36154 MW; 8A2ABD61D2E9816C CRC64;

Query Match 12.6%; Score 73.5; DB 4; Length 357;
Best Local Similarity 25.5%; Pred. No. 8.7;
Matches 35; Conservative 12; Mismatches 35; Indels 55; Gaps 7;

Query Match 12.6%; Score 73.5; DB 4; Length 357;
Best Local Similarity 25.5%; Pred. No. 8.7;
Matches 35; Conservative 12; Mismatches 35; Indels 55; Gaps 7;

RESULT 13
Q9MC06 PRELIMINARY; PRT; 135 AA.
ID Q9MC06;
AC Q9MC06;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF25.
GN ORF25.
OS Bacteriophage D3.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=0042341; PubMed=10572124;
RA Gilkranz Z.A., Kropinski A.M.;
RT "Cloning and analysis of the capsid morphogenesis genes of Pseudomonas aeruginosa bacteriophage D3: another example of protein chain mail?";
RL J. Bacteriol. 181:7221-7227(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20485557; PubMed=11029426;
RA Kropinski A.M.;
RT "Sequence of the Genome of the Temperate, Serotype-Converting, Pseudomonas aeruginosa Bacteriophage D3.;"
RL J. Bacteriol. 182:6066-6074 (2000).
DR EMBL: AF165214; AAP80785.1; -.
SQ SEQUENCE 135 AA; 14876 MW; B50C5ABEAD558EA4 CRC64;

Query Match 12.5%; Score 73; DB 9; Length 135;
Best Local Similarity 27.8%; Pred. No. 3.6;
Matches 20; Conservative 7; Mismatches 31; Indels 14; Gaps 2;
DE Basic helix-loop-helix protein class B 1 (Fragment).
GN BHBL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TAXID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20205664; PubMed=10737801;
RA Wang J., Janni-Sait S.N., Escalon E.A., Carroll A.J., de Jong P.J.,
RA Kirsch I.R., Aplan P.D.;
RT "The t(14;21)(q11.2;q22) chromosomal translocation associated with T-cell acute lymphoblastic leukemia activates the BHBL1 gene.;"
RN Proc. Natl. Acad. Sci. U.S.A. 97:3497-3502(2000).
DR EMBL: AF21520; AFM61215.1; -.
DR InterPro: IPR00102; HHL_basic.
DR InterPro: IPR000408; Reg_chr_condens.
DR Pfam: PF00010; HHL; 1.
DR SMART: SM00355; HHL; 1.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE: PS00626; RCC1_2; UNKNOWN_1.
FT NON_TER 1
FT SEQUENCE 357 AA; 36154 MW; 8A2ABD61D2E9816C CRC64;

RESULT 14
Q9WRN7 PRELIMINARY; PRT; 364 AA.
ID Q9WRN7;
AC Q9WRN7;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE R13 (VIRF).
DR Macaca mulatta rhadinovirus 17577, and
OS Macaca mulatta rhadinovirus 26-95.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
NCBI_TAXID=83534; 119193;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Macaca mulatta rhadinovirus 17577;
MEDLINE=95174001; PubMed=10074154;
RA Seattles P. Bergquam E.P., Axthelm M.K., Wong S.W.;
RT "Sequence and genomic analysis of a Rhesus macaque rhadinovirus with similarity to Kapozi's sarcoma-associated herpesvirus/human
RT herpesvirus 8";
J. Virol. 73:3040-3053(1999).

RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Macaca mulatta rhadinovirus 26-95;
 RC STRAIN-MACACA MULATTA RHADNOVIRUS ISOLATE 26-95;
 RX MEDLINE-20173730; PubMed=10708456;
 RA Alexander L., Denekamp J., Knapp A., Auerbach M.R., Damania B.,
 RA DeGrosiers R.C.;
 RT "The primary sequence of rhesus monkey rhadinovirus isolate 26-95:
 RT sequence similarities to Kapoor's sarcoma-associated herpesvirus and
 RT rhesus monkey rhadinovirus isolate 17577.";
 RL J. Virol. 74:3388-3390(2000).
 DR EMBL; AF210726; AAF60043.1; -.
 DR InterPro; IPR001346; IRF.
 DR Pfam; PF00605; IRF; 1.
 DR SMART; SM00348; IRF; 1.
 SQ SEQUENCE 364 AA; 41189 MW; D8A48195C39DD018 CRC64;
 Query Match 12.5%; Score 72.5; DB 12; Length 364;
 Best Local Similarity 35.1%; Pred. No. 11; Gaps 2;
 Matches 20; Conservative 5; Mismatches 25; Indels 7; Gaps 2;
 Qy 32 TVAGVMSATPKHGLEQCPPAPPVAVTGTGDSAKETVSDQDKRSQGHMWTALPHP 88
 Db 125 TVLNYLEAACVHGLEPGTLPPLPPRPA--EADGARSVYARA-----LATVAPPHP 174
 RESULT 15
 023691 PRELIMINARY; PRT; 554 AA.
 ID 023691;
 AC 023691;
 DT 01-JAN-1998 (TREMBREL. 05, Created)
 DT 01-JAN-1998 (TREMBREL. 05, Last sequence update)
 DT 01-DEC-2001 (TREMBREL. 19, Last annotation update)
 DE Hypothetical protein 60.9 kDa protein.
 GN TIGB16_24.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viriplantae; Streptophytina; Embryophytina; Tracheophytina;
 OC Spermatophytina; Magnoliophytina; eudicots; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis;
 OX NCBI_TAXID:3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NCV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Mitti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
 RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome 1 BAC TIGB16 genomic sequence.";
 RL Submitted (Oct 2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U95973; ARB65493.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 554 AA; 60902 MW; 8188B41F9AE27BBA CRC64;
 Query Match 12.5%; Score 72.5; DB 10; Length 554;
 Best Local Similarity 32.4%; Pred. No. 17; Gaps 1;
 Matches 22; Conservative 3; Mismatches 30; Indels 13; Gaps 1;
 Qy 21 LHPALSLDLYVTVAGVMSATPKHGLEQCPP-----APPPAVTGTGDSAKE 67
 Db 232 LPPPPPLPMAVRKGVAAAPPLPPGTTAALPPPLPMAAGKGVAAPPPLPPGRRGLGAKK 291
 Qy 68 TVSQDKRS 75
 Db 292 VVSKLRS 299

Search completed: June 17, 2003, 10:36:55
 Job time : 84 secs